Inês Isabel Gomes Ferreira

Licenciada em Bioquímica

Study of glycosidic changes in lung cancer: potential effectors in hematogenous metastasis

Dissertação para obtenção do Grau de Mestre em

Bioquímica para a Saúde

Orientador: Doutora Paula Alexandra Quintela Videira, Professora Auxiliar, Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa

Co-orientador: Doutor António Bugalho, Professor Auxiliar, Faculdade de Ciências Médicas, Universidade Nova de Lisboa

Setembro 2015

Inês Isabel Gomes Ferreira

Licenciada em Bioquímica

Study of glycosidic changes in lung cancer: potential effectors in hematogenous metastasis

Dissertação para obtenção do Grau de Mestre em

Bioquímica para a Saúde

Orientador: Doutora Paula Alexandra Quintela Videira, Professora Auxiliar, Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa

Co-orientador: Doutor António Bugalho, Professor Auxiliar, Faculdade de Ciências Médicas, Universidade Nova de Lisboa

Júri: A definir

Nova Medical School - Faculdade de Ciências Médicas da Universidade Nova de Lisboa

Setembro 2015

The work developed until the present date has originated:

- One Poster:

I. Ferreira; M. Carrascal; A. Bugalho; A. Fernandes; P. Baptista; P. Videira. *Role of fucosyltransferase 3 in hematogenous metastasis of lung cancer*. iMED 7.0 Conference Lisbon 2015, Portugal.

Acknowledgements

- Ao grupo de Glicoimunologia do Centro de Estudos de Doenças Crónicas (CEDOC) da Nova Medical School, por me receber no seu laboratório;

- À Professora Doutora Paula Videira, a minha orientadora, pela oportunidade oferecida, orientação prestada e por todo o apoio e ensinamentos durante esta dissertação. Ao Professor Doutor António Bugalho, o meu co-orientador, pela disponibilidade e por toda a ajuda prestada, principalmente a nível clínico;

- À Mylène Carrascal pela amizade, ajuda constante, paciência e disponibilidade para me guiar no decorrer desta tese. Um agradecimento especial por me ter ensinado a gostar "tanto" de fazer Western Blotting's. Sem ela, tudo teria sido mais difícil;

- Ao Doutor Dário Ligeiro por me ter recebido no seu laboratório de Genética Molecular do Centro de Histocompatibilidade do Sul e pela preciosa ajuda em todo o processo de extracção de RNA das amostras e no processo de Real-Time PCR;

- À Graça Marques, Inês Iria, Leonor Rodrigues e Lili pela amizade, conversas, boa disposição, bons momentos e ajuda prestada. Um obrigado especial à Lili pela sua tentativa falhada em tentar cuidar da minha saúde e por me acordar logo de manhã com luz no laboratório;

- À minha italiana preferida Silvia Achilli pela amizade e todos os bons momentos proporcionados durante a sua estadia em Lisboa;

- A todos os restantes membros do grupo de Glicoimunologia por todos os preciosos conselhos, bons momentos e ajuda prestada no decorrer deste trabalho;

- A todos os amigos que me apoiaram no decorrer desta dissertação, em especial à minha Rainha Susy pela amizade, ajuda, apoio, conversas e por ser das pessoas mais incríveis que já conheci. À Princesa Rita por apesar de longe continuar a ser uma grande amiga e um apoio fundamental durante todo este ano. À minha afilhada Raquel por ser a melhor afilhada de sempre e por ser um Exemplo de força e coragem;

- Ao meu melhor amigo e companheiro de vida Luís Garcez por me aturar durante estes últimos anos, apoiar em todos os momentos e nunca deixar de acreditar em mim;

- À minha família. Em especial aos meus pais e irmã, por todo o apoio incondicional, paciência e incentivo prestado. À bebé Constança por ter vindo alegrar estes últimos meses da redacção da tese.

Abstract

Lung cancer (LC), one of the major causes of mortality related to cancer in Portugal, may lead to hematogenous metastasis. Adhesion of cancer cells to endothelium is considered one of the crucial steps involved in metastasis. In blood cells, this adhesion is initiated by endothelial selectin ligands (E-SL) that are glycoproteins or glycolipids decorated mostly with sialyl-Lewis x (sLex) and sialyl-Lewis a (sLea).

While LC has been described as expressing these sialyl Lewis antigens, its functional role in allowing LC adhesion to endothelium is still poorly understood.

We analyzed paired tumor and normal tissues samples from non-small cell lung cancer (NSCLC) patients and three LC cell lines. Immunoblotting assays with anti-sLex/sLea and E-selectin chimera demonstrated that LC tumor tissues significantly overexpress E-SL and flow cytometry results indicated that E-SL are also abundantly expressed in LC cell lines.

To understand the mechanism behind the overexpression of E-SL in LC tissues and cell lines, we analyzed the expression of genes involved in its biosynthesis, namely FUT3, FUT4, FUT6, FUT7, ST3GAL3, ST3GAL4, ST3GAL6, β4GALT1, GCNT1 and GALNT3. It was observed the overexpression of fucosyltransferases FUT3, FUT6 and FUT7 in LC tumor tissues and FUT3 in LC cell lines, being this last one correlated with an increased reactivity of the LC cells to endothelial selectins. It was described that low expression of FUT4 in tumor tissues is correlated with early stages of NSCLC. We also analyzed scaffolds proteins of sLex/sLea and it was identified the carcinoembryonic antigen as an E-SL in NSCLC.

In summary, this thesis contributed to a better understanding of the glycosidic changes and molecules that can influence tumor progression of LC, allowing identifying in the future new diagnosis/prognosis biomarkers and potential therapeutic targets for NSCLC.

Resumo

O cancro do pulmão (LC), uma das principais causas de mortalidade relacionada com cancro em Portugal, pode levar à formação de metástases hematogénicas. A adesão das células tumorais ao endotélio é considerada um dos passos fundamentais envolvidos na metástase. Em células sanguíneas, esta adesão é mediada por ligandos de E-selectina (E-SL), glicoproteínas ou glicolípidos decorados principalmente com sialyl-Lewis x (sLex) e sialyl-Lewis a (sLea).

Tem sido descrito a expressão destes antigénios em LC, contudo o seu papel funcional em permitir a adesão das células de LC ao endotélio é ainda pouco compreendido.

Foram analisadas amostras emparelhadas normais e tumorais de pacientes com cancro de pulmão de não-pequenas células (NSCLC) e três linhas celulares de LC. *Immunoblotting assays* com anti-sLex/sLea e molécula quimérica de E-selectina demonstraram que tecidos tumorais de LC sobreexpressam significativamente E-SL e resultados de citometria de fluxo demonstraram uma expressão elevada de E-SL nas linhas celulares.

Para compreender o mecanismo da sobreexpressão de E-SL em tecidos tumorais e linhas celulares de LC, foi analisada a expressão de genes envolvidos na biossíntese de E-SL, nomeadamente FUT3, FUT4, FUT6, FUT7, ST3GAL3, ST3GAL4, ST3GAL6, β4GALT1, GCNT1 e GALNT3. Observou-se a sobreexpressão das fucosiltransferases FUT3, FUT6 e FUT7 em tecidos tumorais de LC e FUT3 em linhas celulares de LC, sendo que neste último, esta expressão é correlacionada com um aumento da adesão das células de LC às selectinas endoteliais. Foi observado que uma baixa expressão de FUT4 em tecidos tumorais está associada com estadios menos avançados de NSCLC. Foram analisadas ainda proteínas decoradas com sLex/sLea, tendo-se identificado como E-SL o antigénio carcinoembrionário em NSCLC.

Em resumo, esta tese contribuiu para uma melhor compreensão das alterações glicosídicas e moléculas que podem influenciar a progressão tumoral do LC, podendo permitir identificar futuramente novos biomarcadores de diagnóstico/prognóstico e potenciais alvos terapêuticos para o NSCLC.

Index

Index of figures

Index of tables

Index of appendixes

Abbreviations

AD - Adenocarcinoma

ALK - Anaplastic lymphoma kinase

Asn - Asparagine

β4GALT1 - UDP-Gal:betaGlcNAc beta 1,4- galactosyltransferase, polypeptide 1

CD62E - CD62 antigen-like family member E, also known as E-selectin

CD62L - CD62 antigen-like family member L, also known as L-selectin

CD62P - CD62 antigen-like family member P, also known as P-selectin

CEA - Carcinoembryonic antigen

CLA - Cutaneous Lymphocyte-associated Antigen

CMP - Cytosine Monophosphate

C_t - Threshold cycle

DNA -Deoxyribonucleic acid

EDTA - Ethylenediaminetetraacetic acid

EGF - Epidermal growth factor

EGFR - Epidermal growth factor receptor

E-Ig - Recombinant mouse E-Selectin/CD62E Fc chimera

ER -Endoplasmic reticulum

E-SL - E-selectin ligands

FBS - Fetal Bovine Serum

FSC - Forward scatter

FUT - Fucosyltransferase

FUT3 -Fucosyltransferase 3 (galactoside 3(4)-L-fucosyltransferase, Lewis blood group)

FUT4 -Fucosyltransferase 4 (alpha (1,3) fucosyltransferase, myeloid-specific)

FUT6 -Fucosyltransferase 6 (alpha (1,3) fucosyltransferase)

FUT7 -Fucosyltransferase 7 (alpha (1,3) fucosyltransferase)

GalNAc - N-Acetyl- D- galactosamine

GALNT3 - Polypeptide N-acetylgalactosaminyltransferase 3

GAPDH - Glyceraldehyde-3-phosphate dehydrogenase

GCNT1 -Glucosaminyl (N-acetyl) transferase 1, core 2

GDP - Guanine Diphosphate

GlcNAc - N-Acetyl-D-glucosamine

GT - Glycosyltransferase

HCELL - Hematopoietic Cell E-selectin/L-selectin Ligand

IP - Immunoprecipitation

IP-CEA - CEA immunoprecipitate

LC - Lung cancer

MFI - Median Intensity Fluorescence

mRNA - Messenger RNA

MUC1 - Mucin 1

NSCLC - Non-small cell lung cancer

NST - Nucleotide Sugar Transporter

PNGase F - Peptide-N-glycosidase F

PSGL-1 - P-selectin glycoprotein ligand-1

RNA - Ribonucleic acid

RT - Reverse transcription

RT-PCR - Real-time polymerase chain reaction

SCC - Squamous cell lung cancer

SCLC - Small cell lung cancer

SDS-PAGE - Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis

Ser - Serine

Sia - Sialic acid

sLea -Sialyl-Lewis a

sLex -Sialyl-Lewis x

SSC - Side scatter

ST - Sialyltransferase

ST3GAL3 - ST3 beta-galactoside alpha-2,3-sialyltransferase 3

ST3GAL4 - ST3 beta-galactoside alpha-2,3-sialyltransferase 4

ST3GAL6 - ST3 beta-galactoside alpha-2,3-sialyltransferase 6

Thr - Threonine

TNM - Tumor-node-metastasis

TTF-1 - Thyroid transcription factor 1

UDP - Uridine Diphosphate

DISSERTATION

1. Introduction

1.1 Carbohydrates

Carbohydrates are one of the four main classes of organic molecules in living systems, besides the lipids, nucleic acids and proteins (Ghazarian, Idoni and Oppenheimer, 2011). The basic structural units of carbohydrates are monosaccharides. There are nine monosaccharides commonly found in mammals (Figure 1.1).

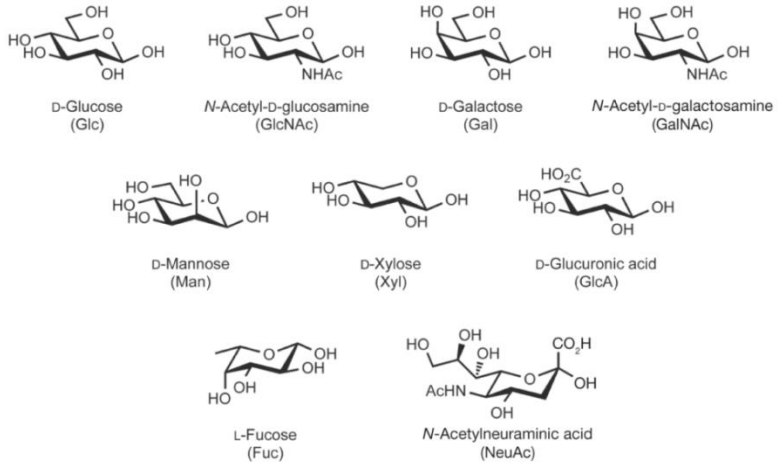


Figure 1.1: Common monosaccharides found in mammals. The abbreviation of monosaccharides is represented in brackets. N-acetylneuraminic acid is also known as sialic acid (Sia). Adapted from Varki *et al.*, 2009.

Monosaccharides are linked together to form oligosaccharides (2-20 monosaccharides residues linked by glycosidic linkages) or polysaccharides (long chains of monosaccharides linked by glycosidic linkages). In most cases, people use the term "glycan" instead of carbohydrate to refer any form of mono-, oligo- or polysaccharide free or covalently attached to another molecule, forming a glycoconjugate.

The glycome, analogous to genome, transcriptome and proteome, involves all the glycans that can be synthesized by an organism. Different types of cells, tissues and organisms, express different types of glycans and the glycome also varies during stages of development/differentiation and during malign transformation (Tateno *et al.*, 2007). The glycome is estimated to be 10-10^4 times bigger than the proteome, depending on species (Freeze, 2006).

The structural diversity of glycans is enormous and results of several carbohydrate characteristics: the number of available monosaccharides building blocks, glycosidic bond position, anomeric configuration (α or β glycosidic linkage), the capacity of glycans to form branched or linear scores structures and carbohydrates modifications (Ghazarian, Idoni and Oppenheimer, 2011).

The glycan structures are not encoded directly in the genome and for that reason, they are called secondary gene products. There are several genes in the genome that after transcription and translation, generate transporters and enzymes (like glycosidases and glycosyltransferases) responsible for the biosynthesis and the assembly of glycans (Varki *et al.*, 2009; Taylor and Drickamer, 2011) .

Glycans are involved in the regulation of multiple cellular mechanisms like: differentiation, cell adhesion, contact inhibition, receptor activation, endocytosis, cell-cell recognition, cell growth and development, anticoagulation, host immune response, metastasis, molecular trafficking and clearance, signal transduction, membrane rigidity, host-pathogen interaction during infection and disease development (Varki *et al.*, 2009; Ghazarian, Idoni and Oppenheimer, 2011; Ohtsubo and Marth, 2006; Raman *et al.*, 2005). Glycans can also affect the intrinsic properties of proteins which they are linked to, such as their solubility in water, their proper protein folding and functional group orientation and protection from proteases (Dall'Olio, 1996).

1.1.1 Glycosylation & Glycoproteins

Glycosylation is a process of the enzymatic addition of glycans to a noncarbohydrate moiety, such as proteins, lipids or other organic compound. It is the most complex form of post-translational modification of proteins. It is important not confuse glycosylation with glycation which is a nonenzymatic and irreversible process of adding glycans, that is elevated in several diseases (Ohtsubo and Marth, 2006).

A glycoprotein is a protein that carries one or more glycans covalently linked to its polypeptide chain. The most common classes of glycoproteins are *N*-glycans and *O*-glycans (Figure 1.2). Glycosylated proteins can be found in all living organisms including eubacteria, *archae* and eukaryotes (Spiro, 2002). The surface of eukaryotic cells is decorated with an array of glycoproteins, forming the glycocalix. Nowadays, it is known that protein glycosylation occurs also in cytoplasmic and nuclear proteins

(Hadley *et al.*, 2014). Glycoproteins are also found at the extracellular matrix that surrounds cells and can be secreted into biological fluids, for example the serum (Taylor and Drickamer, 2011).

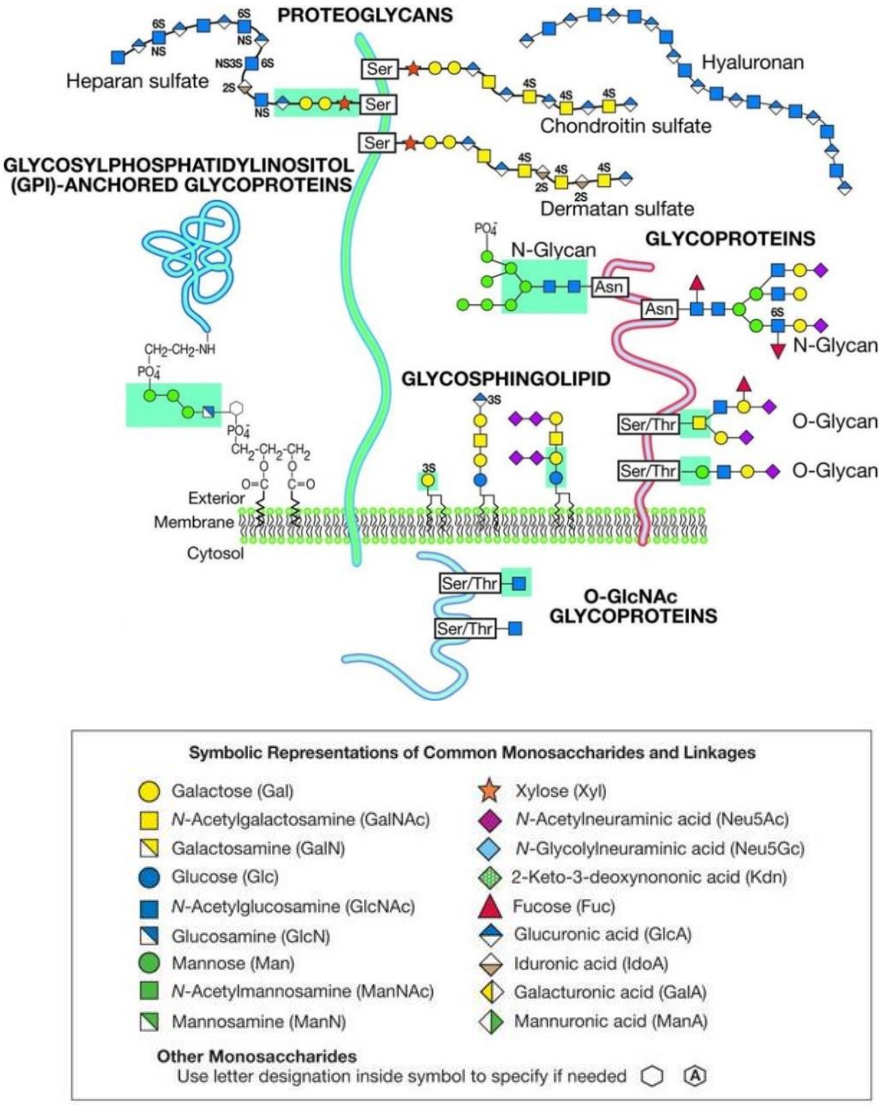


Figure 1.2: Common classes of animal glycans. The monosaccharides are represented in symbols, according to the correspondence shown below. Adapted from Varki *et al.*, 2009.

The biosynthesis of glycoproteins requires activation of monosaccharides to nucleotide-sugars donors. First of all, monosaccharides are obtained by the cell from dietary sources, salvage processes from glycoconjugates degraded within cells and endogenous conversion from other monosaccharides (Yarema and Bertozzi, 2001). The conversion of monosaccharides into nucleotide-sugars donors involves normally the phosphorylation of one or more hydroxyl groups of the monosaccharides. The

nucleotide-sugar donor contains the energy required for the transference of a monosaccharide residue from a nucleotide-sugar donor to the hydroxyl group of an acceptor substrate (in this case, oligosaccharide, monosaccharide or protein), by action of a large group of enzymes called glycosyltransferases (GTs) (Palcic, 1994). GTs are specific for a nucleotide-sugar donor but may recognize more than one different acceptor (Varki *et al.*, 2009). GTs act sequentially, so the product of one GT may yields an acceptor substrate for the action of another GT. They are classified according to the sugar they transfer (Table 1.1).

$$\text{Acceptor} + \text{Glycosyl Donor} \xrightarrow{\text{Glycosyltransferase}} \text{Glycosylated acceptor} + \text{Nucleotide}$$

Table 1.1: The most common glycosyltransferases and nucleotide donors in animal cells. UDP - Uridine Diphosphate; CMP - Cytosine Monophosphate; GDP - Guanine Diphosphate.

Monosaccharide	Glycosyltransferase family	Nucleotide donor
Glucose	Glucosyltransferase	UDP
Galactose	Galactosyltransferase	UDP
N-acetylglucosamine	N-acetylglucosaminyltransferase	UDP
N-acetylgalactosamine	N-acetylgalactosaminyltransferase	UDP
Xylose	Xylosyltransferase	UDP
Glucuronic acid	Glucuroniltransferase	UDP
Sialic acid	Sialyltransferase	CMP
Mannose	Manosyltransferase	GDP
Fucose	Fucosyltransferase	GDP

Besides GTs, glycosidases are also involved in the biosynthesis of glycans. These enzymes are responsible for the hydrolysis of glycosidic linkages to remove sugars from proteins and play an important role in glycan processing in the Golgi apparatus and endoplasmic reticulum (ER), as well as in the modulation of the glycan landscape in the extracellular milieu.

In eukaryotes, most glycosylation reactions occur predominantly in Golgi apparatus and ER lumen. On the other hand, nucleotide-sugars are synthetized in cytosol with exception for CMP-sialic acid synthetized in nucleus (Hadley *et al.*, 2014).

For this reason, the nucleotide-sugars must be transported into the ER lumen or Golgi. This transport is performed by proteins known as nucleotide sugar transporters (NSTs). NSTs are very hydrophobic proteins that function as antiporters: they exchange cytosolic nucleotide sugars into the lumen of the organelles for the corresponding luminal nucleoside monophosphate, such as CMP for CMP-sugars, UMP for UDP-sugars and GMP for GDP-sugars (Hadley *et al.*, 2014).

As previously mentioned, the most common types of glycoproteins are *N*-glycans and *O*-glycans, which will be described below.

1.1.1.1 *N*-glycans

In eukaryotic organisms, *N*-glycans are the most studied form of glycoproteins (90% of glycoproteins are *N*-glycosylated).

N-glycans are covalently attached to proteins at amide nitrogen of asparagine (Asn) side chains. There is a minimal consensus sequence that can accept an *N*-glycan: Asn-X-Ser/Thr (starts with Asn, followed by any amino acid except proline (Pro) and ends with serine (Ser) or threonine (Thr)). The most common sugar linked to Asn is N-acetylglucosamine (GlcNAc), with a β configuration (GlcNAcβ1-Asn).

N-glycans biosynthesis (figure 1.3) begins with the formation of a lipid-linked oligosaccharide (fourteen sugars covalently attached to a lipid dolichol) on the cytoplasmic and luminal face of ER. After that, the entire oligosaccharide is transferred *en bloc* to the Asn residue of a nascent polypeptide by a multi-enzyme complex named oligosacharyltransferase (OST). Following the transference of oligosaccharide, the *N*-linked glycan is processed by glycosidases and GTs, initially in ER and then in Golgi apparatus (Aebi, 2013; Varki *et al.*, 2009; Moremen, Tiemeyer and Nairn, 2012). Processing steps include removal and the addition of sugar residues to form branches or terminal structures.

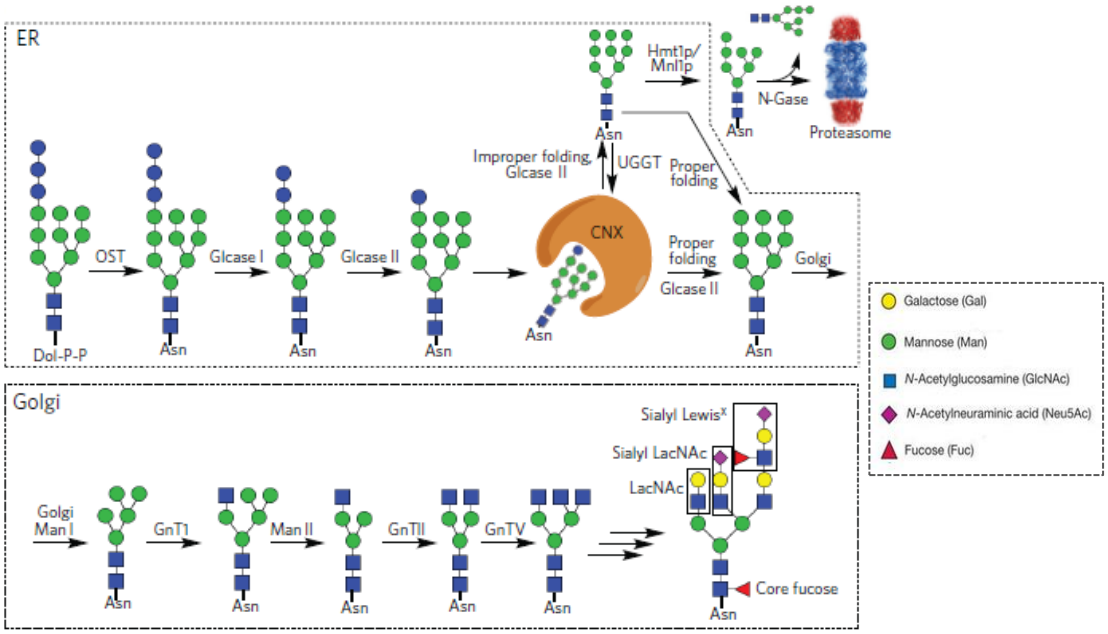


Figure 1.3: Biosynthesis of *N*-glycans. OST, oligosaccharyltransferase; Glcase, glucosidase; UGGT, UDP–glucose-glycoprotein glucosyltransferase; Hmt1p, HnRNP methyltransferase 1; Mnl1p, mannosidase-like protein 1; N-Gase, N-glucosidase; Man, mannosidase; GnT, N-acetyl glucosaminyltransferase; CNX, calnexin. Adapted from Wolfert and Boons, 2013.

1.1.1.2 *O*-glycans

Several types of *O*-glycans have been identified but the most common form is the mucin-type *O*-linked glycans, also called *O*-GalNAc glycans. Mucins are heavily *O*-glycosylated glycoproteins (molecular weight > 200kDa) that can be soluble, secreted or expressed in the membrane. They are present at many epithelial surfaces such as respiratory, reproductive and gastro-intestinal tracts and play an important role in the protection against pathogens.

O-glycans biosynthesis begins in *cis* Golgi apparatus with the binding of N-GalNAc residue to the hydroxyl group of Ser or Thr of a target protein by a polypeptide N-acetyl-galactosaminyltransferase (ppGalNAcT), forming the Tn (Thomsen-nouvelle) antigen. After that, this antigen may be modified or extended with other carbohydrates residues such as galactose, fucose, sialic acid and GlcNAc by specific GTs, which results in many heterogeneous structures (Brockhausen, 1999; Varki *et al.*, 2009; Tian and Hagen, 2009). There are eight *O*-GalNAc glycan core structures, the most common ones are core 1 till core 4 (figure 1.4).

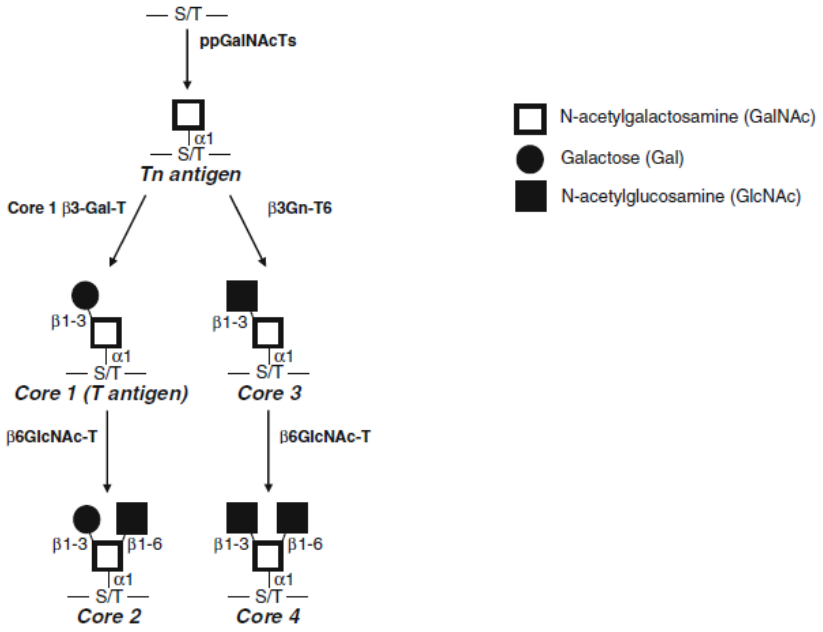


Figure 1.4: Common *O*-GalNAc glycan core structures biosynthetic pathways. ppGalNAcTs, UDP-N-acetylgalactosamine: polypeptide N-acetylgalactosaminyltransferases; Core 1 β3-Gal-T, Core 1 β1-3-galactosyltransferase; β3Gn-T6, β1-3 N-acetylglucosaminyltransferase; β6GlcNAc-Ts, β1-6 N-acetylglucosaminyltransferase). Adapted from Tian and Hagen, 2009.

1.1.2 Altered glycosylation in cancer

Alterations in glycosylation are one of the main characteristics associated with malignant transformation and tumor progression. Cancer progression involves many steps summarize as follows: alterations in extracellular and intercellular signaling pathways, cell growth, detachment of the tumor cells from the primary tumor into the circulation, extravasation of surrounding tissues, cell proliferation and colonization at distant sites (metastasis) (Häuselmann and Borsig, 2014; Reymond, D'Água and Ridley, 2013).

In cancer cells, changes in glycans can take a variety of forms: loss or excessive expression of certain structures, incomplete or truncated glycans and less frequent, the appearance of new antigens. Typically, there are modifications in the expression levels of GTs that lead to the expression of those altered glycans. Besides GTs, modifications in other molecules like transporters and sugar-nucleotides can also alter the phenotype of cancer cells (Kobata and Amano, 2005).

Altered glycosylation in *N*-glycans is normally associated with an increase in β1,6-branching caused by the enhanced expression of β1–6-N-acetylglucosaminyltransferase-5 (GlcNAcT-V) (Dennis, Granovsky and Warren, 1999;

Varki *et al.*, 2009). In cancers with epithelial origin, mucins are the major carriers of incomplete glycosylation in the *O*-linked pathways, that results in the expression of antigens in cell's surface such as sialyl-Tn (sTn, addiction of sialic acid to Tn antigen), sialyl-T (sT, addiction of sialic acid to T antigen), Tn and T antigens, considered tumor-associated carbohydrate antigens (TACA) (Tarp and Clausen, 2008).

This master thesis focuses on one of the most important families of tumor antigens, which include sialyl-Lewis x (sLex) and sialyl-Lewis a (sLea) glycans.

1.1.2.1 Sialyl-Lewis x and a antigens

Lewis epitopes are synthesized by a series of GTs including N-acetilglucosaminyltransferases, galactosyltransferases, sialyltransferases (STs) and fucosyltransferases (FUTs). The sLea and sLex antigens result from mono-fucosylated substitution of the α-2,3- linked sialic acid containing type 1 or type 2 chains (Dall'Olio *et al.*, 2012; Dall'Olio *et al.*, 2014), respectively (Figure 1.5). These two structures are found at terminal ends mainly on the β-1,6 branching of *N*-glycans or *O*-linked glycans attached to glycoproteins and sometimes on glycolipids. Besides cell surface glycoconjugates, these antigens can also be detected in serum.

Clinical studies show an overexpression of sLex and sLea antigens on the surface of tumor cells including colon, gastric, pancreatic and lung cancer (Häuselmann and Borsig, 2014; Mizuguchi *et al.*, 2007; Nakamori *et al.*, 1993; Tozawa *et al.*, 2005) and sometimes their overexpression is correlated with a poor prognosis. Nowadays, sLea is used as a tumor biomarker in pancreatic and colon cancer (Ugorski and Laskowska, 2002).

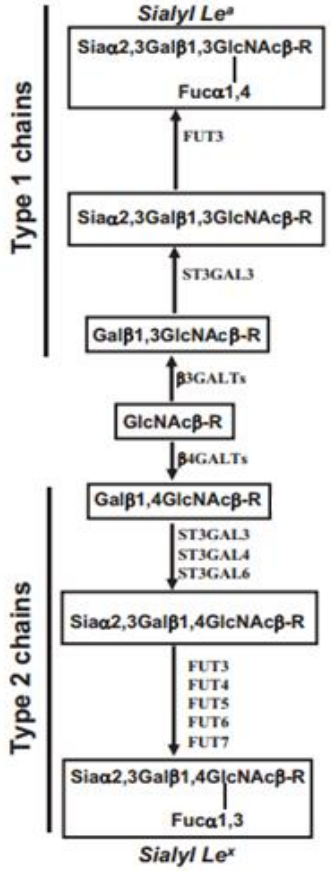


Figure 1.5: Structures and enzymes involved on biosynthesis of sialyl-Lewis x and sialyl-Lewis a antigens. In the figure, sialyl-Lewis x is denominated by Sialyl Le^x and sialyl-Lewis a is denominated by Sialyl Le^a. Adapted from Dall'Olio *et al.*, 2014.

The terminal steps of the biosynthesis of sLex and sLea involve α 1→ 3/4 - fucosyltransferases and α 2→ 3 - sialyltransferases.

FUTs catalyze the transfer of a fucose residue from GDP-α-L-fucose to an acceptor, normally galactose or GlcNAc. This family is composed by three main sub-families: α 1→ 2 FUT, α 1→ 3 FUT and α 1→ 6 FUT. 1→ 3 FUTs (composed by FUT3, FUT4, FUT5, FUT6, FUT7 and FUT9) represent the most important and critical enzymes involved in the last step of the biosynthesis of sialyl Lewis antigens (Chen, 2011).

Typically, one GT only produces one specific glycosidic linkage but FUT3 can produce two glycosidic linkages: α 1→ 3 in sLex and α 1→ 4 in sLea. Importantly, FUT3 is the only enzyme able to synthetize sLea epitope (Dall'Olio *et al.*, 2012).

Although FUT4 can synthetize sLex antigen with a weak activity, it is generally involved in the formation of nonsialyl Lewis antigens, as is FUT9 (Mollicone, Cailleau

and Oriol, 1995). FUT6 is an important enzyme involved in the synthesis of type 2 chain structures and FUT7 is highly expressed on leukocytes and is responsible for the sLex increase in leukemia cell surfaces and other types of tumors (Chen, 2011).

Related to ST family, ST3GALs transfer sialyl group from CMP-Sia to galactose, forming an α 2→ 3 linkage in the final product. ST3GAL family is divided into 6 subtypes. ST3GAL3 uses as substrate type 1 and type 2 chains although it prefers the former. ST3GAL4 and ST3GAL6 prefer type 2 chain as substrate (Chen, 2011).

Increased levels of FUTs and STs have been found in various types of tumors and this expression is correlated with the level of sLex and sLea antigens (Carvalho *et al.*, 2010; Pérez-Garay *et al.*, 2013; Vajaria *et al.*, 2014) .

1.1.2.1.1 Role in metastasis

Tumor cells mimic the process of extravasation used mainly by leukocytes that migrate from the blood to reach sites of injury, infection or inflammation (Strell and Entschladen, 2008). There are four main steps in the extravasation process: tethering and rolling, integrin activation, firm adhesion and transendothelial migration (also called diapedesis). Tethering is characterized by the first contact of the leukocytes with activated endothelium, resulting in rolling of the cells along the surface of the endothelium. Tethering and rolling of leukocytes are mediated by selectins and their glycoprotein ligands. After that, rolling leukocytes in response to specific chemokines and cytokines activate their integrins that leads to a firm adhesion to the endothelium and a following transendothelial migration to the injury site (Ley *et al.*, 2007; Reymond, D'Água and Ridley, 2013; Strell and Entschladen, 2008).

The process of cancer metastasis mimics leukocytes migration and it is facilitated by interactions between tumor cells and endothelial cells in distant tissues, corresponding to the first step of tethering and rolling in leukocytes migration. These interactions are mediated by selectins, a family of calcium-dependent type I transmembrane glycoproteins with a N-terminal C-type lectin domain, an Epidermal Growth Factor (EGF) domain, a variable number of consensus repeats (CR), a transmembrane domain and a short cytoplasmic tail. There are three members in this family: P-selectin (CD62P), L-selectin (CD62L) and E-selectin (CD62E or ELAM-1). P-selectin is expressed by activated endothelial cells and activated platelets, where is

stored in Weibel-Palade bodies and α-granules, respectively. In response to inflammatory cytokines, P-selectin is translocated to cell surface in endothelial cells. L-selectin is expressed in the most of leukocytes and E-selectin is expressed by activated endothelial cells.

The minimal recognition motif for all three selectins is mainly the sLex and/or sLea antigens, so the interaction between cancer cells and selectins is possible because of the presence of these sialylated and fucosylated structures known as selectin ligands on the cell surface of tumor cells (figure 1.6) (Barthel *et al.*, 2007; Gout, Tremblay and Huot, 2008; Läubli and Borsig, 2010) . Of all selectins, E-selectin (107 - 115kDa cell surface glycoprotein) is the major receptor expressed on activated endothelial cells that are involved in adhesion events during metastasis although P-selectin and L-selectin can also contribute for the process (Gout, Tremblay and Huot, 2008; Kannagi *et al.*, 2004). The presence of E-selectin ligands (E-SL) on tumor cells is correlated with high adhesion to activated endothelium. Several E-SL carriers have been identified on tumor cells such as P-selectin glycoprotein ligand-1 (PSGL-1), a glycoform of CD44 known as hematopoietic cell E-selectin/L-selectin (HCELL), several mucins, dead receptor-3 and CD24 among others (Barthel *et al.*, 2007; Burdick *et al.*, 2012; Häuselmann and Borsig, 2014; Läubli and Borsig, 2010).

Higher expression of E-SL in cancer cells is correlated with an enhanced metastatic activity of the cells and is associated with a poor prognosis (Kannagi *et al.*, 2004; Läubli and Borsig, 2010). Furthermore, E-selectin expression is increased on the surface of endothelial cells at proximal sites or directly on tumor metastasis, thus demonstrating a synergy between E-selectin and E-SL to stimulate tumor cell extravasation.

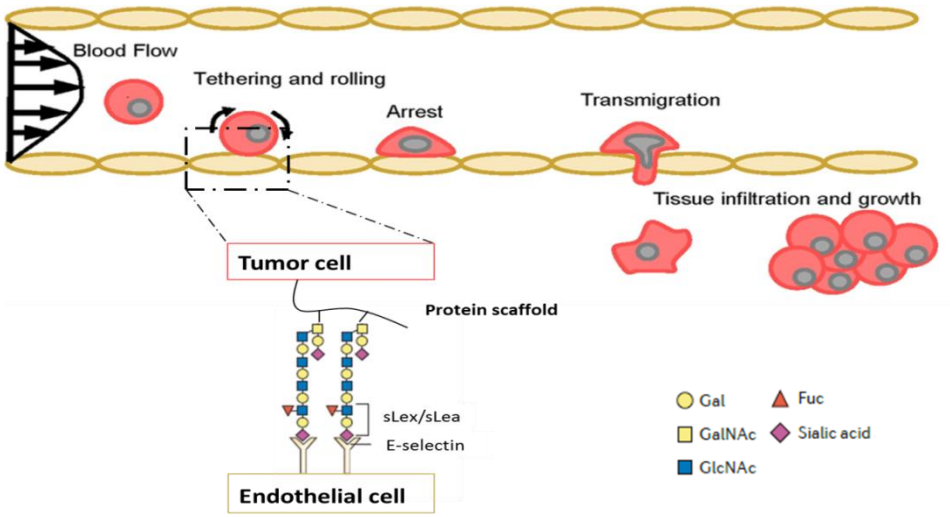


Figure 1.6: Schematic representation of the multi-step process of metastasis in cancer. Adapted from Pinho and Reis, 2015 and Burdick *et al.*, 2012.

1.2 Lung cancer

Lung cancer (LC) is one of the most common forms of cancer (approximately 1.6 million new cases of cancer annually) and the leading cause of death related to cancer in the world, with highest overall prevalence in industrialized regions, such as Europe and North America (Alberg, Brock and Samet, 2005; Cruz, Tanoue and Matthay, 2011). In Portugal it is responsible for 4.200 deaths/year (first cause of cancer related mortality). It occurs predominately in persons between 50-70 years and is more common in men than in women. The global situation is not encouraging, with an overall 5-years survival rate estimated at only 17% in developing countries (American Cancer Society, 2014).

It is characterized by the growth of abnormal cells that starts in the lungs. Over time, these cells continue to grow and divide, forming a mass of tissue called tumor. LC cells can spread to other parts of the body, through blood vessels or lymph vessels and can form new tumors in other tissues (metastasis).

Historically, it was divided in two major types of LC: Small cell lung cancer (SCLC) and Non-small cell lung cancer (NSCLC).

SCLC, as the name indicates, is characterized by small cells that normally spread quickly through the body. It represents 15% of all LC cases, almost always

caused by smoking and typically starts in the bronchi close to the center of the chest. The prognosis of this type of cancer is very poor and there are only two stages in SCLC: limited (tumor only on one side of the chest) and extensive stage (Kitamura *et al.*, 2008).

1.2.1 Non-small cell lung cancer

It is the most common type of LC comprising 85% of all LC cases and is divided in three major subtypes: adenocarcinoma (AD), squamous cell carcinoma (SCC) and large cell carcinoma. These cancer cells tend to grow and spread more slowly than SCLC.

AD represents 40% of LC cases and usually occurs in a peripheral location within the lung. This subtype of cancer arises from mucus-producing cells in bronchial mucosal glands and it is the most common type of cancer in non-smokers. Analysis by immunohistochemistry shows 89% positive for thyroid transcription factor (TTF-1), being a useful marker in the diagnosis of tumors with lung origin (Teh and Belcher, 2014).

SCC usually occurs in the central parts of the lungs, representing 25% of LC cases. It develops in squamous cells (flat cells that line the airways). These cells have an irregular surface, characterized by the presence of keratin pearls. Unlike AD, immunohistochemistry of SCC shows negative for TTF-1 (Teh and Belcher, 2014).

Finally, large cell carcinoma (10% of LC cases) manifests as a peripheral lesion that has a high tendency to metastasize. The remaining 10% of the NSCLC cases correspond to other subtypes much less common or cases when a more specific diagnosis cannot be made (Teh and Belcher, 2014).

1.2.1.1 Etiology

Smoking is the main cause of LC (90% of the patients) with at least 50 carcinogens identified in tobacco smoke. Those components lead to DNA damages, causing activation of proliferation, mutations in tumor suppressor genes such as retinoblastoma and p53, oncogene mutation and/or activation, evasion of apoptosis, angiogenesis and suppression of immune response. One of the most important abnormalities involves mutation in *Ras* family of oncogenes. For non-smokers, a

secondhand smoke (passive smoke inhalation) can increase the risk of developing LC by 30% (Cruz, Tanoue and Matthay, 2011; Groot and Munden, 2012).

Other features can also play an important role in LC such as inherited factors (responsible for 8% of LC cases), atmospheric pollution and exposure to asbestos, radon gas or other chemicals. People with non-malignant lung diseases have a high risk of developing LC.

1.2.1.2 Diagnosis and staging

The symptoms of NSCLC may take a long time to appear and in general, LC is often confused with other less serious conditions. Normally, this type of cancer only produces symptoms in an advanced stage of the disease, so more than 50% of the patients are diagnosed at an advanced stage, which leads to a poor prognosis and survival.

After a comprehensive history and physical exam, a chest x-ray or computed tomography are initially performed. To confirm a diagnosis of cancer, the cells or tissues have to be sampled, normally by minimally invasive methods such as bronchoscopy, transthoracic needle biopsy or thoracentesis (depending on the primary lesion location and potential metastasis).

NSCLC staging uses the TNM (tumor-node-metastasis) classification from American Joint Cancer Committee 7th edition of TNM staging based on International Association for the Study of Lung Cancer (IASLC) (Mirsadraee, 2012). This classification is based on the size of the main tumor (T), the spread of tumor near lymph nodes (N) and presence of metastasis (M). The information of T, N and M categories are combined to assign an overall stage of I, II, III or IV, a process denominated stage grouping. Stages I, II and III are also sub-divided in A or B, according to TNM descriptors such as size, endobronchial location, local invasion among other cancer features. NSCLC patients require an accurate and complete staging because this is an important factor to determine the potential treatment as well as the prognosis. The TNM staging of LC is described in appendix 6.1.

1.2.1.3 Metastasis in NSCLC

Metastatic LC occurs when tumor cells escape from the lung and travel to other parts of the body through blood (hematogenous metastasis) or lymph system. It can be metastatic at the time of diagnosis or during/following the treatment.

LC is the most common cause of brain metastases. Up to 33% of patients with NSCLC develop symptomatic brain metastases during the disease. Besides the brain, LC cells frequently metastasize to the other lung, adrenal glands, bone and liver (Riihimäki *et al.*, 2014; Stenbygaard *et al.*, 1999).

Patients in stage IV are often diagnosed with multiple disseminated metastases (Schuchert e Luketich, 2003).

Metastatic patterns varied depending on histological subtype, gender and age at the time of diagnosis of the disease, however it was reported that SCC is less predisposed to metastasize than AD and the last one prefers spreading to the bone. Survival in metastatic LC is worst in patients with liver or bone metastases (Riihimäki *et al.*, 2014).

According to literature, there were alterations in some tumor suppressor proteins such as p16 that occurs only in metastatic NSCLC and not in primary NSCLC. Besides that, mutations on oncogenes also affect the pattern of LC spread. NSCLC patients with mutations in anaplastic lymphoma kinase (ALK) were more likely to have pericardial or pleural metastases whereas mutations in ALK and epidermal growth factor receptor (EGFR) were more likely to develop liver metastases (Aisner and Marshall, 2012; Doebele *et al.*, 2012).

1.2.1.4 Treatment

Overall, there are three main treatments available for NSCLC that can be used separately or in combination: surgery, radiotherapy and chemotherapy. Depending on the stage of disease, the patients are treated in different ways.

Surgery is the first choice treatment for early-stages, offering a great chance of cure. There are different types of surgery like lobectomy (removing a portion of the lung), pneumonectomy (removing the entire lung) and wedge resection (removing part of pulmonary lobe). Radiotherapy is also an option for early-stages of the disease when the surgery is not possible.

Chemotherapy is characterized by anti-cancer drugs that are delivery into the body. This therapy is used in almost 80% of the patients at some point of the disease: before surgery as a neo-adjuvant therapy, after surgery as adjuvant therapy trying to kill any tumor cells that may have been left behind and for more advanced stages as the main treatment. There are several drugs used in NSCLC treatment but American Society for Clinical Oncology (ASCO) recommends a platinum combination for first-line treatment (Reck *et al.*, 2013).

For advanced stages (stage III and IV), tumors are rarely cured and the main aims are improve and extend the quality of life.

Nowadays, clinical trials are focused on targeted therapy, blocking specific genes or proteins that contribute to cancer survival. This strategy may reduce treatment side effects and is less harmful to healthy cells. For example, AD patients have EGFR mutations in 25% of the cases. Mutations in this receptor result in tumor cells that proliferate uncontrollably. In this case, a target therapy can be used including EGFR inhibition mainly by whole ErbB family blockers (Langer, 2004). Target therapy against angiogenesis is another strategy to stop cancer. Angiogenesis which is characterized by the development of new blood vessels used by tumor cells depends on the interaction of growth factors like vascular endothelial growth factor (VEGF) or fibroblast growth factor (FGF) and respective receptors (VEGFR and FGFR). Inhibition of these molecules may offer a direct anti-tumor effect (Alberg *et al.*, 2005).

1.3 Context and objectives of the work developed for this thesis

Currently, LC is the major cause of mortality related to cancer in Portugal. The critical pathophysiological pathways to the development of NSCLC are barely known. The high incidence of mortality along with the need for an early diagnosis and a proper staging of the disease, determines the urgency to find mechanisms associated with LC and with the development of LC metastasis. A major clinical hurdle that contributes to LC mortality is the metastatic spread of primary tumor cells. Nevertheless the mechanisms triggering hematogenous metastasis of LC are unclear.

LC is known to have an aberrant glycosylation profile however its pathophysiological role is unclear. It is also known to overexpress sialofucosylated glycans, sLex and sLea. These structures depending on their availability at cell surface

have the potential of binding to E-selectin expressed on activated endothelial cells. However, its functional role in allowing LC adhesion to endothelium is still poorly understood.

This thesis had as main aim the identification of new glycan-based biomarkers, in particular sLex and sLea, in NSCLC. It envisaged that a better understanding of their pathophysiological role in tumor cell adhesion to E-selectin, may give us hints to predict tumor progression and metastasis formation. A better understanding of glycosidic changes taking place throughout the tumor progression will help the identification of new diagnostic/prognosis biomarkers and potential therapeutic targets. Thus, this work was outlined in two specific topics:

1- Identification of glycosidic antigens (sLea and sLex) expressed in normal and NSCLC tissues. Normal and tumor matched samples were collected from 48 patients submitted to thoracic surgery based on suspicion of LC.

This topic had the following secondary objectives:

a) Phenotypic analysis of these glycosidic antigens by Dot Blotting to analyze the expression of E-SL in matched normal and tumor LC tissues.

b) Expression analysis of several genes involved in the biosynthesis of sLex and sLea antigens in matched normal and tumor tissues of NSCLC patients by Real-Time Polymerase Chain Reaction (RT-PCR), then deducting the altered biosynthetic pathways. We also analyzed the correlation between the gene expression and the expression of sLex and sLea.

c) Study the role of E-SL expressed at surface of NSCLC tumor cells that contribute to transendothelial migration. During the development of this work, it was investigated the expression of E-SL in tissue samples from NSCLC patients as well as their proteins scaffolds and the pattern of E-SL glycoproteins among the various tumor samples. For these experiences, the main technique used was Western Blotting. A better understanding of the carriers for E-SL will allow the identification of tumor specific targets for novel therapies.

2 - Identification of sLex and sLea in NSCLC cell lines, having the following secondary objectives:

a) Analysis by RT-PCR the expression of genes involved in sLex and sLea biosynthesis in commercial NSCLC cell lines (H292, H1299 and A549).

b) Phenotypic analysis of E-SL by flow cytometry.

c) Functional analysis of E-SL to understand the ability to bind to E-selectin under flow conditions, using adhesion assays.

2. Material and methods

2.1 Culture and maintenance of cell lines

For this work, were used three cell lines derived from non-small cell lung cancer (A549, H1299 and H292) kindly provided by professor Pi-Wang Cheng, from University of Nebraska Medical Center, United States of America. We also used a cell line of colorectal cancer (LS 174T) provided by Professor Fabio Dall'Olio from University of Bologna, Italy. All the cell lines are adherent to plastic surfaces charged negatively and grew up in culture flasks in an incubator with a humidified atmosphere of 5% CO_2 at 37ºC.

The cell line A549 (ATCC® CLL-185™) is derived from a 58 years old Caucasian male lung carcinoma. The cell line H292 (ATCC® CRL-1848™) is derived from a cervical node metastasis of a pulmonary mucoepidermoid carcinoma in a 32 years old Negro female. The cell line H1299 (ATCC® CRL-5803™) is derived from a lymph node metastasis of NSCLC in a 43 years old Caucasian male. The cell line LS 174T (ATCC® CL-188™) was established from a Duke's type B adenocarcinoma of colon in a 58 years old Caucasian female.

H1299 and H292 cell lines were cultivated in Roswell Park Memorial Institute medium-1640 (RPMI-1640), A549 in Dulbecco's Modified Eagle's Medium (DMEM) and LS 1474T in Minimum Essential Medium (MEM), all from Sigma-Aldrich. The basal medium were supplemented with 10% of heat inactivated fetal bovine serum (FBS) from Gibco™, 2mM of L-glutamine (Gibco™) and 100µg/mL of Penicillin/Streptomycin (Gibco™). RPMI-1640 medium was also supplemented with 1mM of sodium pyruvate (Gibco™) and 0.1mM of MEM non-essential amino acids solution (Gibco™).

LC cells lines were cultivated in T-25cm^2 flasks and LS 1474T in T-75cm^2 flasks. The culture medium of the cell lines was renovated each two/three days and the cells were normally detached from the flask at a confluence about 70-80% and passed into new flasks, according to the following protocol:

First of all, we removed the medium of the cultures and washed the cells two times with phosphate buffered saline 1× (PBS 1× - see appendix 6.2). Afterward the cells were released from the bottom of the flask with approximately 0.05 mL/cm^2 of a

0.25% Trypsin- Ethylenediaminetetraacetic acid solution (Trypsin-EDTA solution) from Gibco™, acting for 5min in the incubator, at 37ºC. Then we put complete medium into the T-flask with the released cells, to stop the action of Trypsin-EDTA and collected all the content of T-flask to a 15mL falcon. The cell's pellet was obtained by a centrifugation at 1200rpm for 5min and it was ressuspended in new medium (6mL for T-25cm^2 and 15mL for T-75cm^2). Cells were split in new T-flasks, normally LC cell lines in a proportion of 1:3 and the LS 174T cell line 1:4.

These cell lines were also stored in cryogenic tubes at -80ºC and in liquid nitrogen, using a freezing solution constituted by simple RPMI-1640 medium supplemented with 5% of dimethyl sulfoxide (DMSO) from Merck and 20% of FBS.

2.2 Biological samples

2.2.1 Tissues

Samples collected from 48 patients submitted to thoracic surgery based on suspicion of LC were performed at Unidade de Técnicas Invasivas Pneumológicas in Hospital Pulido Valente (Lisboa), under coordination of Dr. António Bugalho.

For each patient, fragments of tumor tissue and normal pulmonary tissue (local as far as possible of the tumor, likely in a different pulmonary lobule) with a size above 0.5 cm^3 were collected in duplicate for sterile cryogenic tubes, immediately placed in liquid nitrogen.

For this master thesis, normal and tumor tissue samples were divided in two pieces of approximately 40mg: one for RNA extraction and another one to be used in blotting techniques.

2.3 Techniques

2.3.1 Real-Time Polymerase Chain Reaction (RT- PCR)

This technique allows following in "real-time" the amplification of a target deoxyribonucleic acid (DNA) sequence through the addiction of a signal (fluorescence) emitted in each cycle of the reaction. Reverse transcription (RT) of total ribonucleic acid (RNA) into complementary DNA (cDNA) coupling with RT-PCR is a more

sensitive, specific and reproducible method widely used to study the gene expression levels, since combine the process of amplification and detection into a single step.

In my thesis, the fluorescent technology used was *TaqMan* type probe. This probe consists in a specific sequence of oligonucleotides complementary to our gene of interest, with one molecule reporter fluorescent at one end and one molecule quencher at the other. When the reporter is excited by an appropriate laser, it absorbs the light and emits it at a specific wavelength. While the quencher remains near the reporter, it will absorb the emitted light. During the reaction of RT-PCR, the *TaqMan* probe hybridizes with the target sequence and *Taq* DNA polymerase begins the extension of the primers. This polymerase has exonuclease activity 5'→3', which involves the hydrolysis of *TaqMan* probes. During the extension, the *Taq* enzyme encounters the probe that will be degraded. Degrading the probe, the quencher is no longer in close proximity to the reporter and consequently the light emitted by the reporter can be detected. Thus, in each cycle of PCR, the fluorescence increases exponentially, in a way proportional to the amount of the PCR product formed (Wong and Medrano, 2005).

In this work, the quantification strategy used was the relative quantification based on the expression levels of a target gene compared with a control gene, normally a housekeeping gene which is usually abundant and constitutively expressed. This is an essential step to normalize the amount of DNA in all the samples, errors in reverse transcription and also efficiencies variations.

To calculate the relative changes in gene expression, the method used was $2^{-\Delta\Delta C}{}_t$ method, a mathematical model developed by Livak and Schmittgen (Livak and Schmittgen, 2001; Schmittgen and Livak, 2008), with the following arithmetic formula:

$$\text{Fold change (RQ)} = 2^{-\Delta\Delta C}{}_t$$

, wherein the threshold cycle (C_t) is the cycle number at which the fluorescent signal of the reaction crosses the threshold and ΔC_t corresponds to the difference between the value of C_t for the gene of interest and for the endogenous control. Thus, $\Delta\Delta C_t$ is the variation between the value of ΔC_t in test sample and the value of ΔC_t in calibrator sample: in this work, test samples correspond to tumor samples and calibrator samples to normal samples. For this comparative C_t model, we assume that the efficiency of the

gene of interest and the control gene is approximately equal, and it is approximately 100%. If we do not want to compare two conditions (for example: tumor vs. normal or treated vs. non-treated), the relative expression of the gene of interest can be calculated with the expression $2^{-\Delta C}{}_t \times 1000$, that gives us the number of copies of the interest gene for each 1000 molecules of endogenous control. The mRNA expression was normalized using the geometric mean of the expression of two endogenous controls, beta actin and glyceraldehyde-3-phosphate dehydrogenase (GAPDH).

During this master thesis, RT-PCR was performed in 7500 Fast Real-Time PCR System (Applied Biosystems) and the results were analyzed using Sequence Detection Software, version 1.3.

2.3.2 Blotting techniques: western blotting and dot blotting

Western blotting, also called immunoblotting, is a common analytical technique for detecting proteins in a given sample of tissue homogenate or extract. This method usually consists of five main steps: gel electrophoresis, transfer of proteins to a membrane, blocking, detection and visualization.

First of all, the proteins are separated according to their molecular weight by gel electrophoresis on a Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE) and after that migrated proteins are transferred to a membrane. The membrane is then incubated and blocked to prevent non-specific background binding of the detection antibodies. The blocked membrane is incubated with a primary antibody (1º antibody) that recognizes a specific protein of interest. To remove unbound antibodies, the membrane is washed and after that, it is incubated with a secondary antibody (2º antibody) that recognizes specific portions of the 1º antibody. The 2º antibody is conjugated with a reporter enzyme, normally horseradish peroxidase (HRP) or alkaline phosphatase, able to catalyze a chemical substrate to produce either chemiluminescence light that can be detected with photographic film, color that can be visualized on the membrane or fluorescence detected by a camera that captures a digital image of the western blotting. The thickness of the band is reflective of the amount of antibodies bound, thus the amount of the protein of interest as well.

Dot blotting is also a technique used to detecting proteins but it is different from western blotting because in this case, protein samples are not separated by

electrophoresis. Instead, the lysate that contains the protein of interest is applied directly onto a membrane as a circular dot. This is followed by the steps of blocking, detection and visualization as described previously for western blotting. Dot blotting is a less time consuming method but can only confirm the presence or absence of a protein, there is no information about the size of the protein of interest.

2.3.3 Flow cytometry

This technique allows analyzing the physical and chemical characteristics of particles or cells like relative size, granularity and/or fluorescence intensity. A flow cytometer is constituted mainly by three systems: fluidics, optics and electronics.

When a sample in solution enters into a flow cytometer, the cells are focused into a stream of single cells by the fluidics system. After that, cells enter in the optics system that consists of lasers, lenses that collect the emission light and optical filters which direct the resulting light signals to the appropriate detectors. Thus, each cell passes focused laser beam and the light can be refracted (Side Scatter - SSC) or scattered (Forward Scatter -FSC) in all directions. SSC is indicative of the granularity inside the cells and FSC provides information about cell's size. The detected light signals are converted into electronic signals by photodetectors that receive light pulses and the amplifiers associated to them will convert the received signal into a voltage value. The processors will combine all measurements of a particle, making it an event, which can be processed by the computer - electronic system (Radcliff and Jaroszeski, 1998).

One of the most common ways to identify the expression of a particular surface marker in the cells includes the use of fluorophores-labeled antibodies. When a fluorescent molecule conjugated with a monoclonal or polyclonal antibody is added to the cells, the antibody binds to a specific antigen on the cell surface or inside the cell. After the laser light strikes the fluorophore, a fluorescent signal is emitted and can be detected by the flow cytometer. This signal is measured as Median Fluorescence Intensity (MFI), an estimative of the amount of antibodies that bind specific to the target of interest.

During this master thesis, the flow cytometry was performed in Attune® Acoustic Focusing Cytometer (Applied Biosystems), constituted by one blue laser

(488nm) and one red laser (638nm) able to detect six different fluorescence. Blue laser detects four fluorescence's BL1 (green), BL2 (yellow), BL3 (red) and BL4 (red) and the red laser detects RL1 (far red) and RL2 (near IR red). The results were analyzed using the program FlowJo v10.0.7.

2.3.4 Adhesion Assays

These adhesion assays were performed based on the Stamper- Woodruff assay (Stamper and Woodruff, 1976), an in vitro model to study the interaction between lymphocytes and high-endothelial venules of lymph nodes.

In my work, there were slight modifications of the previous method. A recombinant mouse E-Selectin/CD62E Fc chimera (E-Ig) was immobilized on coated slides and then was blocked to reduce the non-specific interactions. This molecule will recognize all the E-selectin ligands, sialylated and fucosylated molecules which bind to the lectin domain of E-selectin.

2.3.5 Immunoprecipitation

Immunoprecipitation (IP) is one of the most common processes that enable the affinity purification of antigens from a complex mixture on a small scale using a specific antibody.

Normally, the antibody for the protein of interest and sample (usually cell lysate) are incubated, so the antibody can bind to the protein in solution. This step is followed by the addition of affinity beads to capture the antigen-antibody complex. In my master thesis, we used agarose beads, that have attached proteins G, immunoglobulins (Ig) binding proteins recognizing and binding preferentially for the heavy chains on Fc region of the antibodies. After binding the antigen-antibody complex to the beads, non-bound components are washed away and the antigen is eluted with an appropriate elution buffer (typically reducing SDS-PAGE sample buffer) that will disrupt the affinity interactions.

There is an important optional step in IP called pre-clearing. Pre-clearing of the sample helps to reduce non-specific binding proteins/ligands to agarose beads. It is performed on the sample before the binding with antibody.

At the end, the purified antigen obtained by IP can be analyzed by western blotting or enzyme-linked immunosorbent assay (ELISA).

2.4 Glycosidic characterization of lung cancer cells and NSCLC tissues

2.4.1 Analysis of gene expression by RT-PCR

2.4.1.1 Total RNA extraction and reverse transcription to cDNA

Total RNA extraction was performed in tissues samples and three LC cell lines. Before the extraction, tissue samples were cut in small pieces and homogenized in 300µL of Mili-Q water, for 20s inside a sterile 2mL tube in a homogenizer (Heidolph DIAX900). Regarding to LC cell lines, the cells were removed from the bottom of the flask as described in section 2.1 of Material and Methods and for each cell line, a pellet between 1- 3×10^6 cells were stored into a 1.5mL tube at -80ºC, until RNA extraction. Tissue samples homogenates and thawed pellets of cell lines were used to proceed to RNA extraction, following the instructions of the NZY Total RNA Isolation kit from Nzytech.

RNA was quantified by spectrometry using a spectrophotometer (Bio-Rad SmartSpec™ Plus) and the purity of the samples was analyzed by the ratio of absorbance (Abs) at 260nm and 280nm (Abs_{260}/Abs_{280}). The concentration of the samples was determined using the following relation: one unit of Abs at 260nm corresponds to 40µg of RNA/mL in solution.

Conversion to cDNA was performed using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems). A mixture of 50µL was made by reaction containing 3.5µL MultiScribe™ reverse transcriptase (50U/µL), 10µL reverse transcription buffer (10×), 10µL random primers (10×), 4µL deoxynucleotides (dNTP's mix - 25×) and 22.5µL of mili-Q water. This mixture was added to RNA samples in a proportion of 1:1, in PCR microtubes. Reverse transcription was performed in a thermal cycler (PTC-100™ Programmable Thermal Controller) with the following program: 10min at 25ºC (step 1), 120min at 37ºC (step 2) and 5min at 85ºC (step 3). If the samples were not taken off after these steps, they remained inside thermal cycler at 4ºC up to 24h after the step 3. After the conversion, cDNA samples were stored at -20ºC until RT-PCR analysis.

2.4.1.2 Real Time-PCR

Each reaction of PCR was performed in a final volume of 15µL, constituted by 7.5µL of TaqMan® Universal Master Mix II no AmpErase® uracil-N-glycosylase (UNG) 2× (Applied Biosystems), 1.7µL cDNA, 2.8µL mili-Q water and 3µL of Taqman probes and primers (TaqMan® Pré-Developed Assay Reagents 20× - Applied Biosystems) specific for each gene of interest. The genes analyzed by RT-PCR are described in table 2.1. For tissue samples, all the genes in table 2.1 were analyzed except for the CEACAM5 gene (this gene was analyzed only in LC cell lines).

Table 2.1: List of genes analyzed in tissue samples and LC cell lines by RT-PCR: ACTB and GAPDH were used as endogenous controls.

Gene symbol	Gene name	Assay identification*	UniGene ID**
ACTB	Beta Actin	Hs99999903_m1	60
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	Hs99999905_m1	2597
β4GALT1	UDP-Gal:betaGlcNAc beta 1,4-galactosyltransferase, polypeptide 1	Hs00995149_m1	2683
CEACAM5	Carcinoembryonic antigen-related cell adhesion molecule 5	Hs00944025_m1	1048
FUT3	Fucosyltransferase 3 (galactoside 3(4)-L-fucosyltransferase, Lewis blood group)	Hs00356857_m1	2525
FUT4	Fucosyltransferase 4 (alpha (1,3) fucosyltransferase, myeloid-specific)	Hs01106466_s1	2526
FUT6	Fucosyltransferase 6 (alpha (1,3) fucosyltransferase)	Hs00173404_m1	2528
FUT7	Fucosyltransferase 7 (alpha (1,3) fucosyltransferase)	Hs00237083_m1	2529
GALNT3	Polypeptide N-acetylgalactosaminyltransferase 3	Hs00237084_m1	2591
GCNT1	Glucosaminyl (N-acetyl) transferase 1, core 2	Hs00946933_m1	2650
ST3GAL3	ST3 beta-galactoside alpha-2,3-sialyltransferase 3	Hs00196718_m1	6487
ST3GAL4	ST3 beta-galactoside alpha-2,3-sialyltransferase 4	Hs00272170_m1	6484
ST3GAL6	ST3 beta-galactoside alpha-2,3-sialyltransferase 6	Hs00196086_m1	10402

*Assay identification is the identification given for Applied Biosystems that refers to a set of probes and corresponding primers for each gene. The prefix "Hs" is an abbreviation for *Homo sapiens*, specie in which the assay was formulated. The suffix "m1" indicates an assay whose probes and primers were designed across an exon-exon junction and "s1" indicates an assay whose probes and primers were designed within a single exon. ** UniGene ID refers to the identification number for each gene in UniGene NCBI database (http://www.ncbi.nlm.nih.gov/unigene).

The mixture of all the components was made in a 96-well microplate (Fast Optical 96-Well Reaction, Applied Biosystems). After sealed, the plate was centrifuged at 2000rpm during 5min. The thermal cycling conditions were 1 cycle of 2min at 50ºC, 1 cycle of 10min at 95ºC and 50 cycles of 15s at 95ºC and 1min at 60ºC. Each reaction was made in duplicate and to normalize the values of each sample was used the mean of C_t values of two endogenous controls (β-actin and GAPDH).

The analysis of gene expression was performed using $2^{-\Delta\Delta C}{}_t$ method, as explain in topic 2.3.1 of this master thesis.

2.4.2 Phenotypic analysis

2.4.2.1 Expression of sLex and sLea antigens in tissues samples by dot blotting

Almost all the tissue samples analyzed by RT-PCR were also studied by dot blotting, except samples from patient number 15, 20 and 21.

In order to release the proteins of interest, tissues samples needed to be lysed. For each sample, the piece of tissue kept for blotting techniques was cut and placed in a 2mL round bottom microfuge tube and homogenized in 330µl of lysis buffer (see Appendix 6.2) for 20s at homogenizer. Then the homogenates were maintained in constant agitation for 2 days at 4ºC. After the lysis, the samples were centrifuged at 17000xg for 2min and the supernatant was saved and used for all dot blotting and western blotting experiences.

In order to know the amount of protein to place in each dot, lysates were quantified using the protocol recommended by Pierce® BCA Protein Assay Kit (Thermo Scientific).

All the samples were tested in two concentrations: 2µg/5µL (per dot) and 10µg/5µL (per dot). The lysates were diluted in lysis buffer to perform the required final concentration and applied slowly onto the dots created in the membrane of nitrocellulose (Nitrocellulose Blotting Membrane 0,45µm, Amersham™ Protran™), avoiding touching the membrane with the pipette tip. After absorption of the lysates, the membrane was incubated in blocking solution (non-fat milk 10% diluted in PBS 1×-Tween 0.1%) during the night at 4ºC under agitation. Next morning, the blocking solution was removed and the membrane was incubated with HECA-452 antibody

(1:1000; Biolegend) diluted in PBS 1×-Tween 0.1% for 1h at room temperature under agitation. After the incubation, the antibody solution was removed and the membrane washed three times (5min each) with PBS 1×-Tween 0.1%. After that, the membrane was incubated with secondary antibody Anti-Rat IgM-HRP (1:2500; SouthernBiotech) diluted in PBS 1×-Tween 0.1% for 1h at room temperature under agitation. At the end of the incubation period, the antibody solution was removed and the membrane was washed three times (5min each) with PBS 1×-Tween 0.1%. Finally, it was added to the membrane a solution with two reagents of the detection kit Lumi-Light Western Blotting Substrate (Roche) in a ratio of 1:1 in order to cover the entire surface of the membrane for 1min. The excess liquid was removed and the membrane inserted in a plastic sleeve into a *cassette.* In dark room, a photographic film (Amersham Hyperfilms™ ECL) was placed on the membrane for 2min. Then, the film was passed in revelation solution (Sigma-Aldrich), washed in distilled water, passed in fixer solution (Sigma-Aldrich) and washed again in distilled water. Depending on the quality of revelation, the membrane was exposed at different exposure times (2min and 5min). Image's analysis was performed using ImageJ 1.48v software.

2.4.2.2 Expression of E-SL, sLex/sLea antigen, mucin 1 (MUC1), CD44 and carcinoembryonic antigen (CEA) in tissues samples by western blotting

The tumor tissues samples that showed a high expression of sLex and sLea by dot blotting were chosen to be analyzed by western blotting. Also the matched normal tissues samples were analyzed.

Lysate samples were ressuspended in a solution containing loading buffer (5× SDS-PAGE sample loading buffer; Nzytech) with β-mercaptoetanol (Sigma-Aldrich) at a final concentration of 20µg/10µL and boiled during 7min at 95-100ºC in order to denature the proteins.

It was prepared a polyacrylamide gel, specifically a 8% resolving gel and a 4% stacking gel (see constitution of both gels in Appendix 6.2) in Multiple Gel Caster (Amersham Biosciences). After solidification, the gel was clamped into an electrophoresis apparatus with Running Buffer 1× (see appendix 6.2) and samples, positive control and molecular weight markers (NZYColour Protein Marker II;

Nzytech) were loaded onto gel. Initially, the gel was run at low voltage (100V) and after the samples entered on the resolving gel, the voltage was increased to 150V to separate the proteins until the end of the gel. When the dye molecule (bromophenol blue dye present in loading buffer) reached the bottom of gel, the power was turned off.

After that, it was prepared the cassette of the transfer system (Mini-PROTEAN Tetra System; Bio-Rad). First of all, two fiber pads and two filter papers were soaked in transfer buffer 1× (see Appendix 6.2) and then one fiber pad was placed on the gray side of the cassette. Then, a sheet of filter paper was place on the fiber pad and the gel was put gently on it. A membrane of PVDF (Sequi-Blot™ PVDF, membrane roll, 0.2µm, Bio-Rad; activated previously 5min in methanol- VWR) was placed on the gel, being careful to remove any air bubbles which may have formed. The cassette was completed by placing a piece of filter paper on the membrane and the last fiber pad.

The cassette was placed inside the transfer system with a frozen blue cooling unit and filled to the "blotting" mark on the tank with transfer buffer 1×. At the power supply, it was selected a time of 90min and amperage of 400mA.

After completed transfer, power supply was turned off and the membrane incubated in blocking solution (non-fat milk 10% diluted in TBS 1× -Tween 0.1%; see constitution of TBS 1× in Appendix 6.2) during the night at 4ºC under agitation. The following steps (incubation with primary antibody, secondary antibody, detection and revelation) were the same as described for dot blotting experiences, being the membrane incubated with each antibody for 1h at room temperature and washed between them to remove unbound antibodies. All the molecules/antibodies used in western blotting experiences are described in table 2.2.

Table 2.2: List of molecules/antibodies and respective dilutions used in western blotting experiences: * Diluted in TBS 1×-Tween 0.1%; ** Diluted in TBS 1×-Tween 0.1% + 2mM $CaCl_2$.

Antigen	1º Staining (dilution ratio, manufacturer)	2º Staining (dilution ratio, manufacturer)	3º Staining (dilution ratio, manufacturer)
sLex and sLea	HECA-452 (1:1000 *; Biolegend)	Anti-Rat IgM-HRP (1:2500*; SouthernBiotech)	
CD44	Anti CD-44 clone 2C5 (1:1000*; Biolegend)	Anti-mouse Ig-HRP (1:2500*; BD Pharmigen™)	
MUC1	Mucin-1 VU4H5 (1:100*; Santa Cruz)	Anti-mouse Ig-HRP (1:2500*; BD Pharmigen™)	
CEA	Anti-human CD66e (1:1000*; ImmunoTools)	Anti-mouse Ig-HRP (1:2500*; BD Pharmigen™)	
E-SL	Recombinant mouse E-selectin/CD62E Fc Chimera - E-Ig (1:500 **; R&D Systems)	Anti-m CD62E (1:1000 **; BD Pharmigen™)	Anti-Rat IgG HRP (1:2000**; SouthernBiotech)

The washing buffer used was TBS 1× -Tween 0.1%, except for E-SL staining that was TBS 1×-Tween 0.1%. + 2mM $CaCl_2$.

In the case of MUC1 staining, it was used the membrane previously used for sLex/sLea staining. After the final step of revelation in sLex/sLea staining, the membrane was washed two times (10min each) with TBS 1×-Tween 0.1% to remove chemiluminescent substrate. Then, the membrane was incubated in Western Blotting Stripping Buffer (Thermo Scientific) during 15min under agitation to remove rests of primary and secondary antibodies without removing or damaging the immobilized antigen. At the end of incubation, it was washed two times (10min each) with TBS 1×-Tween 0.1%. Finally, the membrane was blocked with non-fat milk 10% diluted in TBS 1×-Tween 0.1% during the night at 4ºC under agitation and the following steps were the same as described previously.

Relatively to E-SL staining, there was an extra step: the addiction of a third antibody. At the end of the incubation period, the secondary antibody solution was removed and the membrane washed three times (5min each) with TBS 1×-Tween 0.1% + 2mM $CaCl_2$. After that, the membrane was incubated with third antibody for 1h at

room temperature under agitation. The next steps were the same as described previously.

Depending on the quality of revelation, the membrane was exposed at different exposure times (30s, 1min, 2min, 10min, 30min).

2.4.2.3 Immunoprecipitation of CEA in tissue samples

First of all, it was added 80µl of beads-protein G in a tube denominated clearing and 60µL in a tube denominated IP. After that, it was added 1mL of lysis buffer to wash the beads and the tubes were inverted and centrifuged at 12000×g for 1min at 4ºC. The beads were washed three more times with 1mL of lysis buffer (centrifuge to pellet the beads between each wash). Supernatant was carefully removed after each wash to not let the beads dry.

To blocking non-specific interactions, each tube was incubated with 500µL of lysis buffer with bovine serum albumin (BSA from Sigma Aldrich - 0.04g for each 4mL of lysis buffer) for 1h at 4ºC under constant agitation. After that the beads were centrifuged at 12000×g for 1min at 4ºC and washed three times with 1mL of lysis buffer (centrifuge to pellet the beads between each wash). The lysate of the sample of interest (100µg/300µL in lysis buffer) was added to the clearing tube and incubated for 2h at 4ºC under constant agitation. With this step (called pre-clearing), the proteins in lysate that bind non-specific to the beads will be removed. After pre-clearing, the lysate was completely removed and transferred to a new 1.5mL tube, where it was added 3µL of anti-human CD66E (1µg/1µL; ImmunoTools) and incubated for 2h at 4ºC under constant agitation. Finally, the mix of lysate and antibody was added to the IP tube and incubated overnight at 4ºC under constant agitation.

Next morning, the tube was centrifuged at 12000×g for 1min at 4ºC and the supernatant was completely removed. Then the beads were washed three times with 1mL of lysis buffer (centrifuge to pellet the beads between each wash). After the last wash, the supernatant was completely aspirated and it was added 30 µL of loading buffer to the bead pellet. The tube was heated at 95-100ºC for 7min to elute the antigen of interest. Finally, the tube was vortexed and centrifuged at 12000×g for 1min at 4ºC to pellet the beads. The supernatant was completely collected, transferred to a clean tube and frozen at -80ºC for later use in western blotting.

Later, it was run one gel with positive controls, molecular weight markers and two lanes (approximately the correspondence to 50µg of total lysate in each lane) of CEA immunoprecipitate (IP-CEA). After the process of transference, the membrane was cut in two pieces and one was stained with HECA-452 and another one with E-Ig. The procedure of western blotting is indicated in previous point of Material and Methods (2.4.2.2).

In order to know the type of glycosylation of sLex/sLea in CEA molecule, it was made the immunoprecipitation of CEA from tumor lysate of the patient that showed the highest expression of CEA and after that, two conditions were tested: immunoprecipitate treated with peptide-N-glycosidase F (PNGase F) or non-treated. Besides CEA molecule, it was also studied the general type of glycosylation of E-SL glycoproteins presented in tumor lysate from this patient, so tumor lysate was treated and non-treated using the same protocol as for IP-CEA.

For the treatment with the enzyme, the protocol used was in denaturing conditions and all the reagents used were from PNGase F kit (New England BioLabs® Inc.). First of all, it was mixed IP-CEA (equivalent to the amount of CEA present in 20µg of total tumor lysate) or 20µg of tumor lysate and 1µL of 10× glycoprotein denaturing buffer to make a 10µL total reaction volume and the mixture was heated at 95ºC for 10min. After that, it was placed in ice and centrifuged for 10s at 12000×g. To make a total reaction volume of 20µL, it was added to previous tube 2µL of 10× GlycoBuffer 2, 2µL 10% NP-40 and 3µL of H_2O. Finally, 3µL of PNGase F was added and the tube was mixed gently and incubated at 37ºC for 6h. In non-treated tubes, the procedure was the same but the enzyme was substituted by water. After incubation period, the tubes were heated at 95ºC for 5min in loading buffer 1× and frozen at -80ºC for later use in western blotting.

Later, it was run one gel with a positive control, molecular weight markers and four lanes corresponding to IP-CEA treated, IP-CEA non-treated, total lysate treated and total lysate non treated. After the process of transference, the membrane was stained with HECA-452 antibody. The procedure of western blotting is indicated in previous point of Material and Methods (2.4.2.2).

2.4.2.4 Cell lines staining with antibodies by flow cytometry

At a confluence about 70-80%, lung cancer cells from H292, H1299 and A549 were detached from the T-flask as described in section 2.1 of Material and Methods. For each cell line, the pellet was ressuspended in 3mL of PBS 1×, the cells were counted with Neubauer Chamber and divided by 12 microfuge tubes of 1.5mL (0.14×10^6 cells/tube). After that, the cells were ressuspended in 100µL PBS 1× and were stained with antibody anti-human CD66e (3µL; ImmunoTools), commercial antibody anti mucin-1 VU4H5 (5µL; Santa Cruz), antibody anti mucin-1 (15µL; hybridoma supernatant from clone HMGF-2), antibody anti mucin-5B (5µL; Santa Cruz), antibody anti mucin-5AC (5µL; Santa Cruz), HECA-452 (3µL; Biolegend) and antibody CD44-APC (5µL; Biolegend) for 30min at 4ºC.

After incubation period, the cells were washed with PBS 1× and stained with Anti-Mouse FITC (1µL; Dako) except for CD44 staining and HECA-452 staining, the last one stained with Anti-Rat Ig APC (2µL; BD Pharmigen™). The incubation period was 15min in the dark at room temperature. For CD44 staining, no secondary antibody was need and after incubation the cells were washed and ressuspended in 1mL of PBS 1× and the tube was kept in dark until the end of all procedure.

The cells were also stained for E-SL. For that, two tubes were needed: one for staining in the presence of EDTA and another one for calcium. In the first tube, the cells were ressuspended in 100µl PBS 1× + EDTA (2mM in Mili-Q water) and were stained with chimera E-Ig (1 µL; 0.5 µg/µL) and Anti-human IgG-FITC (0.75µL; Sigma Aldrich). In second tube, the cells were ressuspended in 100ul PBS 1× + $CaCl_2$ and were stained with same antibodies as the previous tube. The incubation period for both tubes was 1h in the dark at 4ºC.

Three tubes were used as negative controls: only cells and cells staining only with 2º antibody (2ºAb Anti-Mouse FITC and 2ºAb Anti-Rat Ig APC).

After all incubation periods, all cells were washed and ressuspended in 1mL PBS 1×, proceeding to flow cytometry analysis. The results were presented as MFI for each marker to detect. This value was normalized against the MFI value obtained in negative controls: cells stained with 2ºAb Anti-Mouse FITC for MUC1, MUC5AC, MUC5B, and CEA; cells stained with 2ºAb Anti-Rat Ig APC in CD44 and HECA

staining; and for E-SL, cells with PBS 1× + EDTA + chimera E-Ig + Anti-human IgG-FITC.

2.4.3. Functional analysis of E-selectin ligands in LC cell lines by adhesion assay

0.25µg of E-Ig (see appendix 6.2) in 5µL (diluted in mili-Q water) were spot in each coated slide (Thermo Scientific Superfrost® Plus) and allowed to dry. After that, a circular grid was drawn on the slides to indicate the regions that have the adhered protein. In these zones, the slides were blocked with 400µL of 1% BSA in Mili-Q water, for 1h at room temperature. After this period, the excess of liquid was removed and allowed to dry.

At a confluence about 70-80%, LC cells from H292, H1299 and A549 cell lines were detached from the T-flask as described in section 2.1 of Material and Methods. For each cell line, the pellet was ressuspended in 3mL of PBS 1×, the cells were counted with Neubauer Chamber and divided by 2 microfuge tubes of 1.5mL (1×10^6 cells/tube). There were two conditions to test in these assays: cells in the presence of calcium and cells in EDTA. These two conditions are essential because the binding to E-selectin is dependent on the presence of calcium and by other way, EDTA is a chelating agent that chelates calcium molecules. Then, the cells were centrifuged at 1.5×g for 2min and ressuspended in 200µL Hanks' Balanced Salt Solution - modified (HBSS; Sigma-Aldrich) + 5mM $CaCl_2$ or 200µL HBSS + 2mM EDTA.

For each cell line, there were two slides: one for cells in calcium buffer and another one for cells in EDTA buffer.

Then, the cells were put above the E-Ig coated slides and were incubated for 30min at 4ºC under agitation. After this incubation period, the slides were washed by immersion (three times) in HBSS + $CaCl_2$ or HBSS + EDTA to remove the unbound cells and then the remaining cells were fixed with 3% glutaraldehyde (CarlRoth) (diluted in HBSS + $CaCl_2$) for 10min at 4ºC. At the end, the slides were washed in distilled water for 5min and the bound cells were observed in a microscope (Olympus TL4 IXS1). The number of bound cells was counted using ImageJ 1.48v program.

2.5 Statistical analysis

Data from normal tissues were paired with data from tumor tissues and statistical differences were analyzed using Wilcoxon matched-pairs signed rank test. To analyze the differences between fucosyltransferases and sialyltransferases in tumor samples it was performed Dunn's multiple comparison test. The correlation between genes was analyzed using Spearman Correlation. It was used the Mann-Whitney U-test to investigate relations between gene mRNA expression or sLex/sLea expression with patient categories.

All the tests applied were non-parametric tests since the sample size was not large and did not fulfill the requirements for parametric statistical analysis.

Tests were considered statistically significant when $p < 0.05$ (*), $p < 0.01$ (**), $p < 0.001$ (***), $p < 0.0001$ (****) and marginally significant when $0.05 < p < 0.1$. Statistical analysis was performed using GraphPad Prism 6 (GraphPad Software, Inc).

3. Results

3.1. Patient's characteristics

The tissues analyzed in this master thesis were collected from patients with a diagnosis of non-small cell lung cancer. These patients were submitted to surgery and the harvest of the tissues was performed immediately after lobectomy/segmentectomy (less than 10 min), being the samples placed in liquid nitrogen. After collection, samples were sent to the laboratory of Glycoimmunology of the Faculty of Medical Sciences (NOVA University of Lisbon) and stored in liquid nitrogen until further processing. Normal and tumor matched samples from 48 patients were analyzed in this study. The determination of the histological type was performed by the laboratory of Anatomic Pathology from Hospital Pulido Valente in Lisbon.

Thirty-three patients had a positive diagnosis for NSCLC so they were used for RNA extraction. The main characteristics of these patients are summarized in the following table, with the most detailed data in attachment (see appendix 6.3).

Table 3.1: Clinical characteristics of patients enrolled in this study: type of NSCLC, gender, stage and information about smoking. The number of cases (percentage) is represented. * AD corresponds to Adenocarcinoma and SCC to Squamous Cell Carcinoma.

.

Type of NSCLC	**AD*** -29 (87.9%) **SCC*** - 4 (12.1%)
Gender	**Male**- 26 (78.8%) **Female** - 8 (21.2%)
Smoker or Non-smoker	**Smoker** - 25 (75.8%) **Non-smoker** - 9 (24.2%)
Stage	**IA** - 13 (39,4%) **IB** - 4 (12.1%) **IIA** - 1 (3 %) **IIB** - 2 (6.1 %) **IIIA** - 13 (39.4 %)

21 patients were analyzed by RT-PCR. Twelve patients were excluded from this study because it was not possible to extract RNA from tissue with a good quality or

quantity. The main characteristics of these patients are summarized in the following table, with the most detailed data also in appendix 6.3.

Table 3.2: Clinical characteristics of the 21 patients analyzed by RT-PCR: type of NSCLC, gender, stage, age and information about smoking. The number of cases (percentage) is represented.

Type of NSCLS	**AD** -17 (81%) **SCC** - 4 (4%)
Gender	**Male** - 16 (76.2%) **Female** - 5 (23.8%)
Smoker or Non-smoker	**Smoker** - 14 (66.7%) **Non-smoker** - 7 (33.3%)
Stage	**IA** - 7 (33.3%) **IB** - 3 (14.3%) **IIA** - 1 (4.8 %) **IIB** - 1 (4.8 %) **IIIA** - 9 (42.8 %)
Median age and Minimal/maximal age	65-years (Min= 48-years; Max= 83-years)

3.2 Glycosidic characterization of normal and tumor tissues from NSCLC patients

3.2.1 Analysis of gene expression by RT -PCR

Here, we aimed to analyze the gene expression of relevant glycosyltransferases involved in the biosynthesis of sLex and sLea, by RT-PCR, in normal and tumor matched tissues samples from NSCLC patients, then deducting the altered biosynthetic pathways. In addition, if possible, we analyzed the relations between gene expression levels and patient's characteristics.

The selected genes for gene expression analysis were FUT3, FUT4, FUT6, FUT7, that code for fucosyltransferases; ST3GAL3, ST3GAL4, ST3GAL6, that code for sialyltransferases; β4GALT1, GCNT1 and GALNT3. Genes coding for fucosyltransferases and sialyltransferases were chosen because they are key enzymes involved in the last steps of sialyl Lewis antigens biosynthesis. Although FUT5 is also

related with the biosynthesis of sLex and sLea, its expression was not studied since it is not expressed in most tissues *in vivo* (Chen, 2011). β4GALT1 (see figure 1.5 in Introduction), GCNT1 (essential for the formation of core-2 *O*-glycan; see figure 1.4 in Introduction where this gene belongs to the family β6GlcNAc-T) and GALNT3 (responsible for the initiation of mucin-type *O*-glycosylation; see figure 1.4 in Introduction where this gene belongs to the family ppGalNAcTs) are genes involved in early steps of the *N*/*O*-glycosylation and according to the literature, there is an altered expression of those genes in LC (Dosaka-Akita *et al.*, 2002; Machida *et al.*, 2001).

We also used information contained in a database called cBioPortal database for Cancer Genomics, to understand if the selected genes would be a good target for study. This database is an open-access resource that provides analysis of large-scale cancer genomics data sets. The users can compare gene alterations across multiple cancer studies, visualize patterns of gene alterations across samples in a cancer study and the portal gives also information about survival analysis, analysis of mutual exclusivity between genomic modifications and correlations plots (Cerami *et al.*, 2012; Gao *et al.*, 2013). There are four steps to performing a query of a single-cancer study: 1- select a cancer study, 2- select one or more genomic profiles (ex: mutation, mRNA expression), 3- select a patient set and 4 - enter one or more genes that people want to analyze.

In this database, were analyzed two studies about LC: one about adenocarcinoma and another one about squamous cell carcinoma. The collected database data are presented in Appendix 6.4. In both cases, it was observed an up-regulation of selected genes with a higher up-regulation in SCC study (43.8% in SCC vs. 24.3% in AD). Alterations in selected genes in AD seem to decrease month survival of patients although it is not statistically significant. Furthermore, there are several cases of gene pairs with co-occurrent alterations: FUT3-FUT6 in SCC study; FUT4-FUT6, FUT4-GALNT3, FUT3-GALNT3, FUT3-ST3GAL4, FUT3-FUT6, FUT3-FUT4 and B4GALT1-ST3GAL4 in AD study, which means, when one determined gene is altered in one cancer sample, there is also in the same sample an alteration in the other gene pair.

Besides these selected genes playing an important role in the biosynthesis of sialyl Lewis antigens, they also have a high importance in LC, as shown by the above data. Thus we proceeded to the analysis of them in the collected samples by RT-PCR.

The comparison of gene expression values between normal and tumor tissues of NSCLC patients for each gene is represented in figure 3.1.

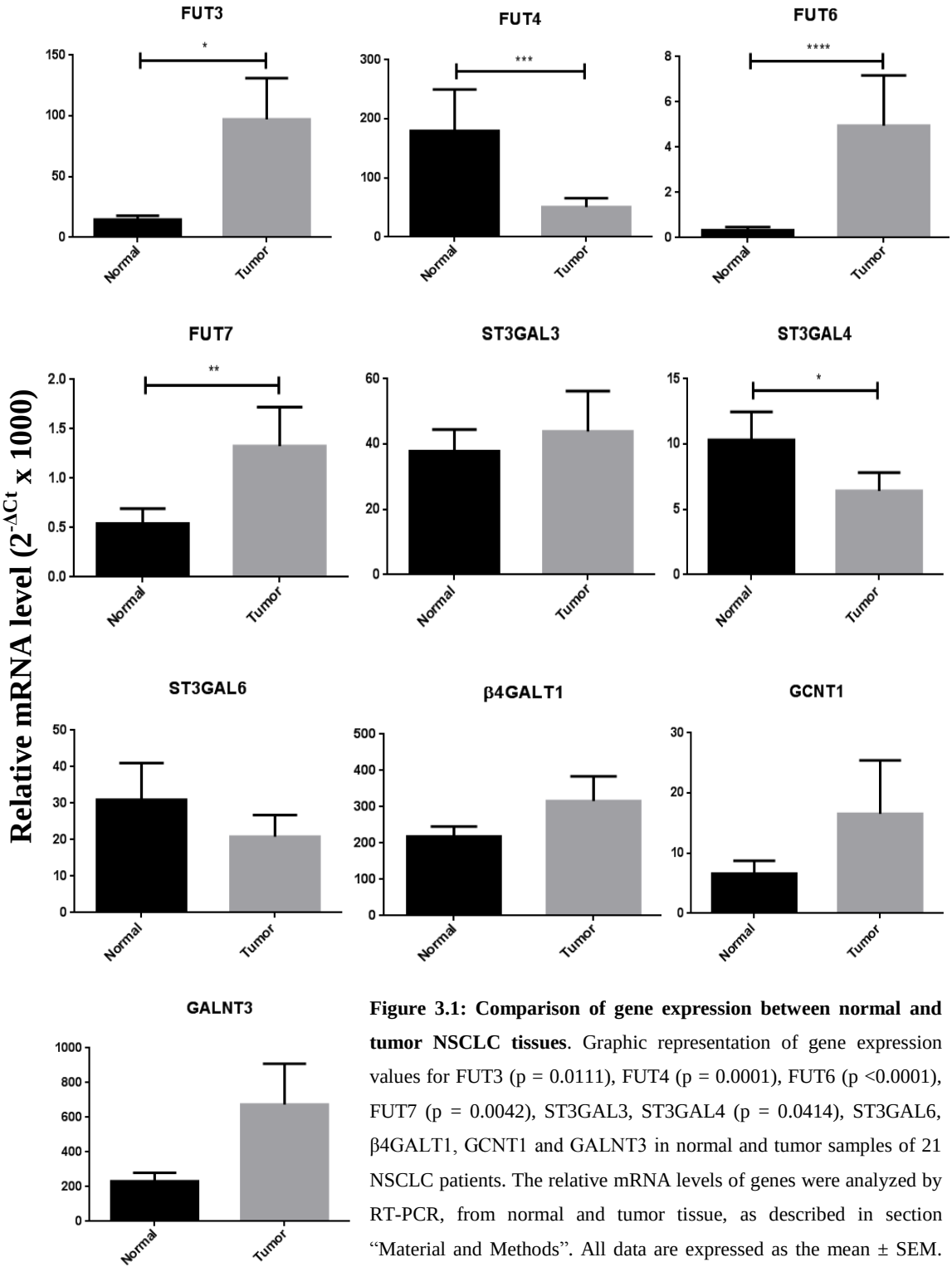


Figure 3.1: Comparison of gene expression between normal and tumor NSCLC tissues. Graphic representation of gene expression values for FUT3 ($p = 0.0111$), FUT4 ($p = 0.0001$), FUT6 ($p < 0.0001$), FUT7 ($p = 0.0042$), ST3GAL3, ST3GAL4 ($p = 0.0414$), ST3GAL6, β4GALT1, GCNT1 and GALNT3 in normal and tumor samples of 21 NSCLC patients. The relative mRNA levels of genes were analyzed by RT-PCR, from normal and tumor tissue, as described in section "Material and Methods". All data are expressed as the mean ± SEM. All p-values were calculated using Wilcoxon matched-pairs signed rank test. Values statistically significant: $p < 0.05$ (*), $p < 0.01$ (**), $p < 0.001$ (***), $p < 0.0001$ (****).

Table 3.3: Fold change expression (RQ) between mRNA expression of analyzed genes in tumor samples and normal samples. Fold change expression was calculated using the formula $2^{-\Delta\Delta Ct}$, as described in the Material and Methods section. RQ value presented for each gene was calculated as the geometric mean of all RQs calculated in all the patients analyzed.

Genes	RQ (×)	Genes	RQ (×)
FUT3	2.92	ST3GAL4	0.50
FUT4	0.36	ST3GAL6	0.51
FUT6	7.64	β4GALT1	1.05
FUT7	2.13	GCNT1	1.30
ST3GAL3	0.63	GALNT3	1.48

The gene expression of FUT3 (p = 0.0111, 2.92 times more express in tumor samples), FUT6 (p <0.0001, 7.64 times more express in tumor samples) and FUT7 (p = 0.0042, 2.13 times more express in tumor samples) are significantly increased in tumor samples compare to normal samples, in opposite to FUT4 (p = 0.0001, 3.6 times more express in normal samples) and ST3GAL4 (p = 0.0414, 5 times more express in normal samples) that are significantly reduced (figure 3.1 and table 3.3). The expression of the other analyzed genes in tumor samples was not significantly different from the correspondent normal samples. In tumor tissues, the most expressed fucosyltransferase is FUT3 with statistically differences with FUT6 (p=0.002) and FUT7 (p <0.0001) as you can see in figure 3.2. In parallel, the most expressed sialyltransferase is ST3GAL3 with statistically differences with ST3GAL4 (p <0.0001).

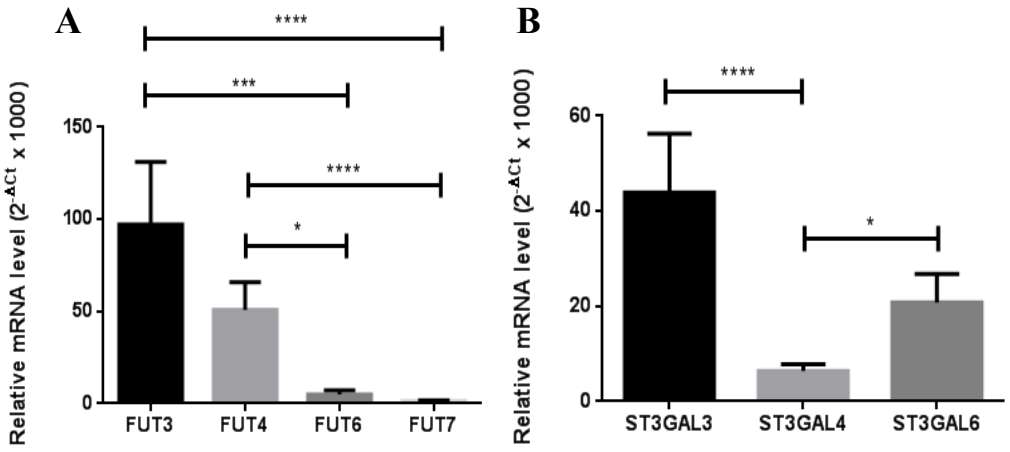


Figure 3.2: Comparison of the expression of fucosyltransferases and sialyltransferasse in lung tumor tissues. Graphic representation of gene expression values in tumor samples of 21 NSCLC patients for fucosyltransferases (A) and sialyltransferases (B)**.** The relative mRNA levels of genes were analyzed by RT-PCR, as described in section "Material and Methods". All data are expressed as the mean ±SEM. All p-values were calculated using Dunn's multiple comparison test. Values statistically significant: p <0.05 (*), p<0.001 (***), p<0.0001 (****).

Once some genes have been found with significant changes between normal and tumor samples from NSCLC patients, it was investigated the relation between mRNA expression of these genes in tumor samples with patients categories: smoker vs. non-smoker, age (less vs. greater than 65 years old - median age) and less advanced stage of the disease (IA + IB) vs. more advanced stage of disease (IIIA). As you can see in figure 3.3, there was no statistical association with smoking and age of the patients for any of the studied genes, except for FUT3 in smoker patients (82.42 ± 44.84) compare with non-smoker patients (126.58 ± 50.93) with p = 0.0823. However, it was possible to see that patients with more of 65-age shown a higher expression of all studied genes compare with the patients less than 65-age.

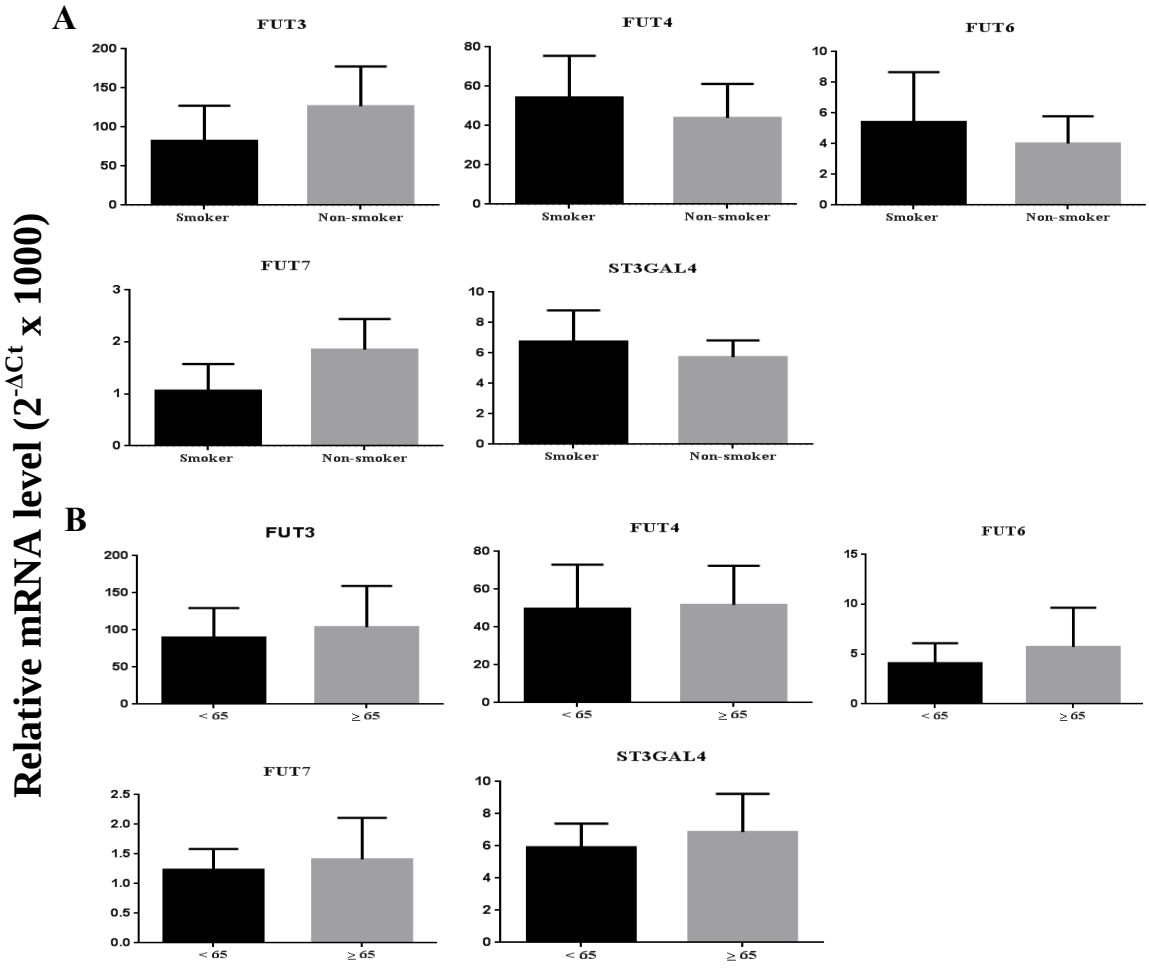


Figure 3.3: Comparison of fucosyltransferases and sialyltransferases expression in tumors, according to smoke habits and age. Graphic representation of gene expression values in tumor samples of 21 NSCLC patients for FUT3, FUT4, FUT6, FUT7 and ST3GAL4. **A**: Smoker patients vs. Non-smoker patients **B:** Patients were divided into two groups based on the values below and above the median age (65). The relative mRNA levels of genes were analyzed by RT-PCR, as described in section "Material and Methods". All data are expressed as the mean ±SEM. All p-values were calculated using Mann-Whitney U-test. Tests were considered marginally significant when $0.05<p<0.1$.

Analyzing figure 3.4, it was found that FUT4 expression was significantly increased (p = 0.0142) in stage IIIA compared with stage I (IA and IB), with 76.59 ± 23.79 (IIIA) and 15.22 ± 8.015 (IA + IB) relative mRNA levels, respectively. The expression of FUT3, FUT6, FUT7 and ST3GAL4 was not statistically different between initial stages and advanced stage.

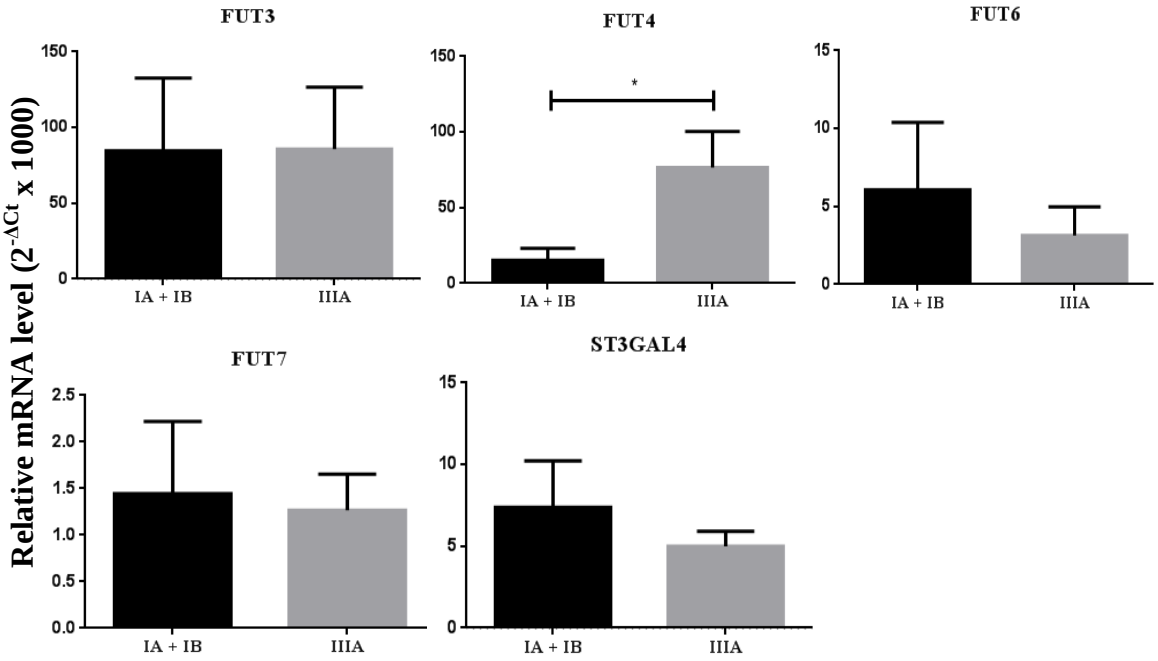


Figure 3.4: Comparison of the fucosyltransferases and sialyltransferases expression in tumors, according to tumor stage. Graphic representation of gene expression values in tumor samples of 21 NSCLC patients for FUT3, FUT4, FUT6, FUT7 and ST3GAL4. Patients were divided into two groups: less advanced stage (stage IA and IB) and more advanced stage (IIIA). The relative mRNA levels of genes were analyzed by RT-PCR, as described in section "Material and Methods". All data are expressed as the mean ±SEM. All p-values were calculated using Mann- Whitney U-test. Values statistically significant: $p < 0.05$ (*).

After comparisons between categories of patients' characteristics, it was verified whether there was any correlation between the ten genes analyzed in tumor and normal samples by the Spearman correlation. These results are represented in appendix 6.5.

In tumor samples, there was a statistically significant positive correlation between FUT3 mRNA expression and all the other genes expression, except for FUT4. FUT6 only had a positive correlation statistically significant with GCNT1 (p=0.037) and FUT3 (p=0.009). FUT7, ST3GAL4, ST3GAL6 and β4GALT1 showed a positive correlation statistically significant with the majority of genes except FUT4 and FUT6. ST3GAL3 and GALNT3 presented a positive correlation statistically significant for all

genes except FUT6. GCNT1 was the only gene with a positive correlation statistically significant for all studied genes. Explaining the results for the previous gene, in tumor LC tissues there is a positive correlation between GCNT1 and all the other genes, meaning that when its expression increases there is also an increased for all the other genes analyzed with all p-values less than 0.05.

In normal samples, there was a negative correlation between FUT3 and FUT4 and also between FUT4 and β4GALT1 but with no statistical significance. FUT6 only had a positive correlation statistically significant with FUT3 ($p = 0.039$). GALNT3 had a positive correlation statistically significant with all genes except FUT6. β4GALT1 only showed a positive correlation statistically significant with FUT3 and GALNT3. FUT7 had a positive correlation statistically significant with ST3GAL3, ST3GAL4, ST3GAL6, GCNT1 and GALNT3. FUT4 had a positive correlation statistically significant with ST3GAL3, ST3GAL4, GCNT1 and GALNT3.

As expected, all the genes described as involved in the biosynthesis of sialyl Lewis antigens are correlated on multiple gene pairs with significant correlations. Overall, comparing correlations results in normal and tumor samples there is a greater correlation between genes in tumor samples (73.3% of correlation - 45 results of which 33 have a significant level less than 5%) than in normal samples (51.1% of correlation - 45 results of which 23 have a significant level less than 5%).

3.2.2 Expression of sLex/sLea antigens in tissues samples by dot blotting

Initially, normal and tumor tissues were lysed and quantified for protein content. The values obtained are presented in Appendix 6.6.

To evaluate the expression of sLex and sLea in normal and tumor tissues, the samples studied by RT-PCR were also analyzed by dot blotting except samples from patient number 15, 20 and 21, because there was no sufficient amount of tissue to make the lysates. In this technique, it was used the antibody HECA-452 to detect specifically sLex and sLea antigens. The monoclonal antibody HECA-452 recognizes a 140kDa homodimer protein named Cutaneous Lymphocyte-associated Antigen (CLA). This antigen is expressed on T cells in skin, Natural Killer cells, memory B cells, dendritic cells, as well as on granulocytes, monocytes and activated endothelial cells. CLA

corresponds to a posttranslational modification of a surface glycoprotein, P-selectin glycoprotein ligand 1 (PSGL-1) that is constituted by a carbohydrate epitope of sialic acid and fucose that made sLex/sLea antigens. CLA is a subset ligand for E-selectin, P-selectin and L-selectin.

The majority of the patients showed a higher expression of sLex and sLea in tumor samples compared with normal samples (figure 3.5). Only patient number 4 and number 11 had a higher expression of sLex and sLea in normal samples compared with tumor samples, as you can see in the values of relative density represented in Table 3.4. In patients 2, 6, 7, 16 and 19 there was at least a two-fold change in expression of antigens between tumor and normal samples.

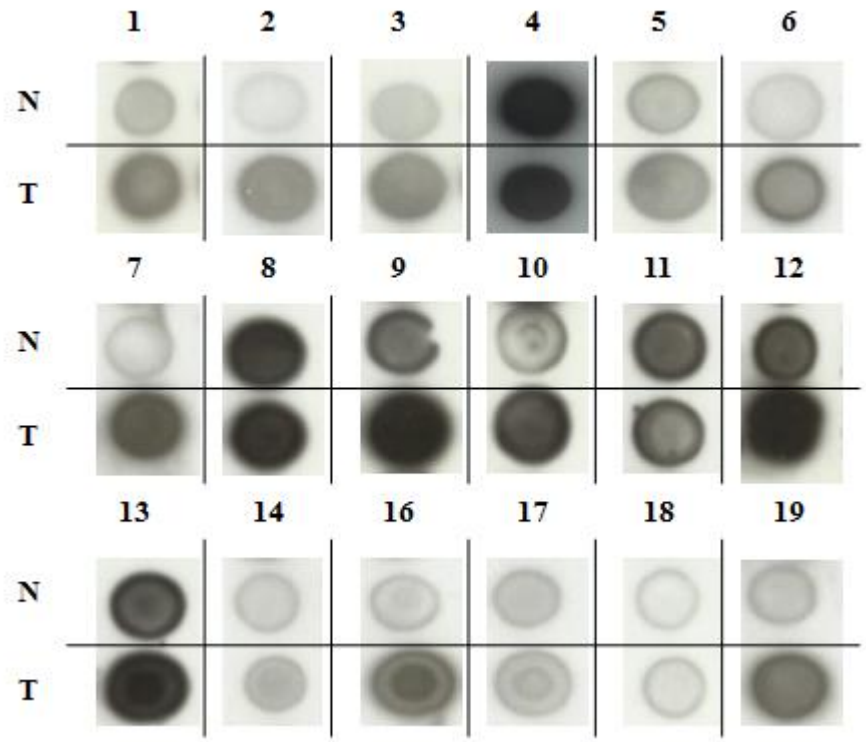


Figure 3.5: Expression of sLex and sLea in NSCLC patient's matched normal/tumor tissue pairs. Matched tumor (T) and normal samples (N) from NSCLC patients were homogenized, dot blotted (10µg/5µL) and stained with antibody HECA-452. The numbers correspond to the patient number. Time of *films* exposition: 2min.

Table 3.4: Values of relative density for each patient. Relative density was calculated by dividing the percent value for each tumor sample by the percent value for the correspondent normal sample. The percent value corresponds to the percentage of the total size of all the measured dots. These results are relative to figure 3.5. This quantification was performed using ImageJ 1.48v *software*.

Patient	Relative density	Patient	Relative density
1	1.77	**10**	1.7
2	2.52	**11**	0.87
3	1.97	**12**	1.51
4	0.88	**13**	1.34
5	1.43	**14**	1.1
6	2.07	**16**	2.68
7	2.64	**17**	1.22
8	1.07	**18**	1.13
9	1.74	**19**	2.09

We then assessed the relation between sLex/sLea expression with gene expression of the patients in tumor samples. Patients were divided into two groups (low and high expression) based on the values below and above median of the sLex/sLea expression in dot blotting. The median value for sLex/sLea expression was calculated based only in the values of sLex/sLea expression obtained for tumor samples analyzed by dot blotting. These values were calculated by ImageJ 1.48v software and they were represented as Area, which indicates the intensity of the blot as a numerical value. The median value was 21436.6145.

It was demonstrated that a higher expression of ST3GAL3 (p-value=0.0502, figure 3.6) and a higher expression GCNT1 (p-value=0.0502, figure 3.6) were associated with a higher expression of sLea and sLex antigens. There was no statistical significant association for any of the other studied genes.

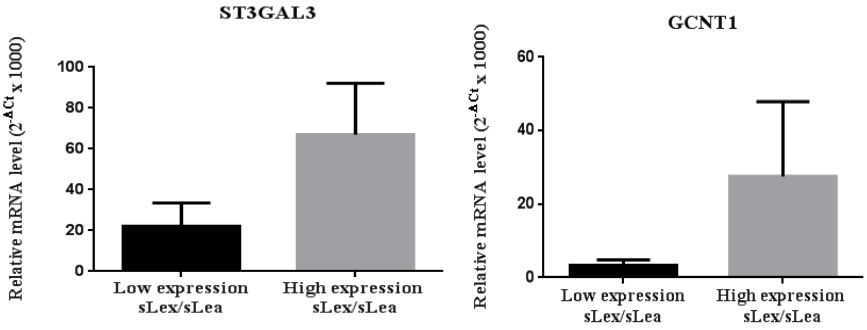


Figure 3.6: Comparison of ST3GAL3 and GCNT1 expression in tumors, according expression of sLex/sLea. Graphic representation of gene expression values in tumor samples of 18 NSCLC patients for ST3GAL3 and GCNT1. Data are parsed in low (n=9) and high expression (n=9) based on the values below and above median of the sLex/sLea expression in dot blotting. The relative mRNA levels of genes were analyzed by RT-PCR, as described in section "Material and Methods". All data are expressed as the mean ±SEM. All p-values were calculated using Mann-Whitney U-test.

It was also assessed the relation between sLex/sLea expression with the stage of disease, gender of the patient and smoking habits of the patients. For each of these studies, there was no statistical significant association.

3.2.3 Expression of sLex/sLea antigens and E-SL in tissues samples by western blotting

Initially, it was analyzed the expression of E-SL glycoproteins and sLex/sLea antigens in tumor samples from NSCLC patients by western blotting technique with two main goals:

- To study the pattern of E-SL glycoproteins and compare it between the various tumor samples;

- To verify if there was E-SL glycoproteins that were not decorated with sLex/sLea, been therefore not recognized by antibody HECA-452, in order to better characterize all E-SL in NSCLC tissues.

To detect specifically sLex and sLea antigens, it was used the monoclonal antibody HECA-452, the same used in dot blotting technique. To recognize all E-SL, it was used the Recombinant Mouse E-selectin/CD62E Fc Chimera (E-Ig) that recognizes sialylated and fucosylated molecules which bind to the lectin domain of E-selectin. It is important to refer that human and mouse E-selectin proteins share 81% amino acid similarity.

Eight patients with high values of relative density were selected for western blotting analysis. The selected patients were patients with the number 1, 2, 3, 5, 6, 7, 9 and 10. After that, four polyacrylamide gels were made: two for HECA-452 staining and two for E-Ig staining. In each gel, one lane was used as positive control normally used by the laboratory group with a sample with a known pattern of protein bands for HECA-452 and E-Ig staining. A positive result from the positive control, even if the samples are negative, will indicate the procedure is working well. In the remaining lanes, it was placed 20µg of each tumor lysate. The membranes were stained with HECA-452 and E-Ig, and after membrane revelation, it was possible to see the pattern of protein bands recognized by each molecule (see next figure: figure 3.7).

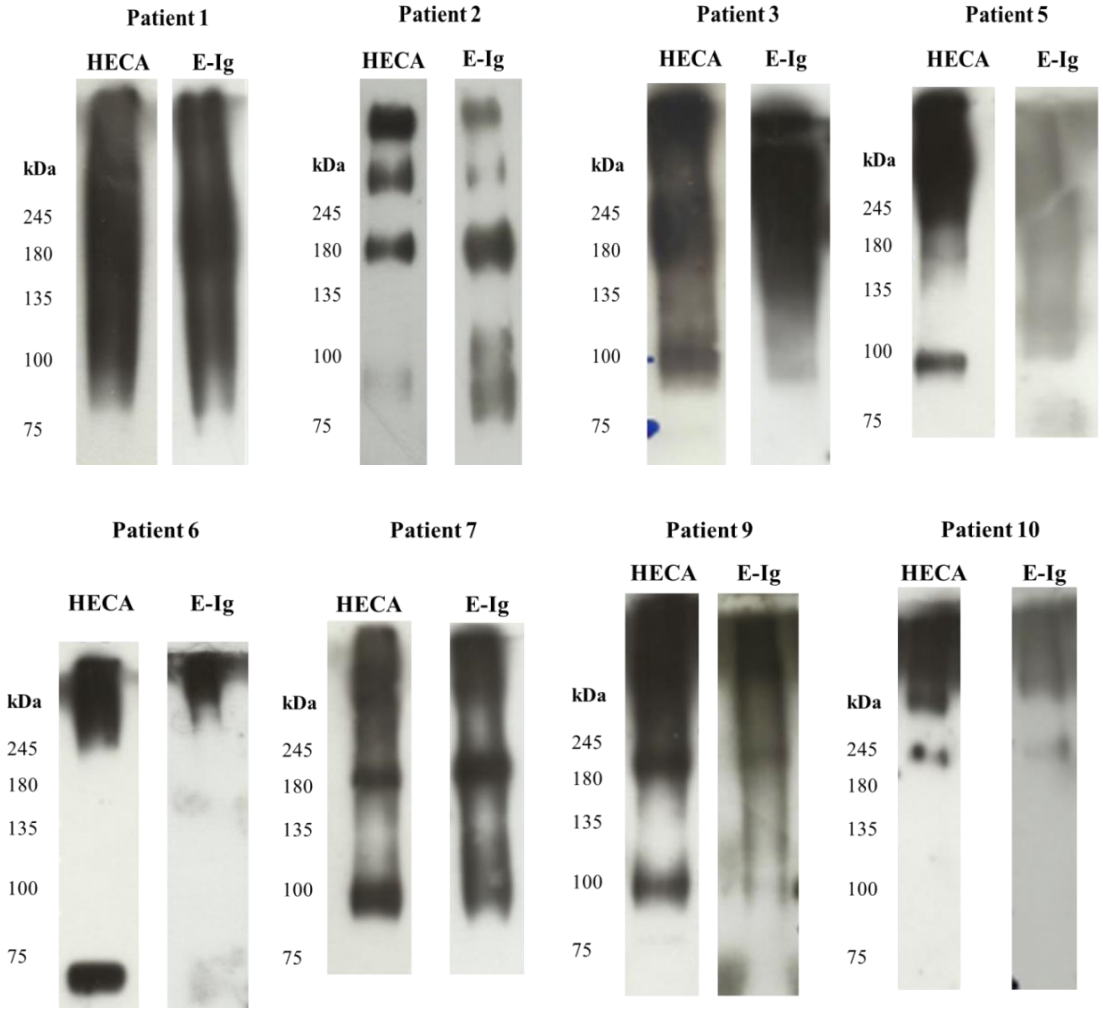


Figure 3.7: Western Blotting analysis of E-SL proteins present in tumor lysates from eight NSCLC patients. The tumor tissue from patient 1, 2, 3, 5 , 6, 7, 9 and 10 was lysed and the western blotting procedure was performed as described in section "Material and Methods". For each patient's tissue, first lane corresponds to anti-sLex/sLea, i.e HECA-452 staining (HECA) and second lane corresponds to chimera E-Ig staining. Band sizes are indicated in kilodaltons (kDa) at the left of each set.

For patient 1, staining of the tumor tissue with HECA-452 and E-Ig was very similar, although it was not possible to observe the presence of a specific band in the two staining's. The labeled glycoproteins containing sLex/sLea and being E-SL were presented in a wide range of molecular weights above 75kDa, being detected a major signal around 180kDa.

The western blotting from tumor tissue of patient 2 revealed four bands in HECA-452 staining (two intense bands above 245kDa, one intense band slightly above 180kDa and one band not too intense near 100kDa) and five bands in E-Ig staining, the same observed with HECA-452 staining and one more band slightly above 100kDa, representing the latter a glycoprotein that is not decorated with the known sLex or sLea glycan structure but it is able to bind to E-selectin.

For patient 3, it was impossible to observe the presence of a specific band in E-Ig staining. However, in this staining it was possible to observe a large area with an intense staining above 180kDa. Relatively to HECA-452 staining, a band at 100kDa and an intense band between 180-245kDa was observed.

In the western blotting's from tumor tissues of patient 5 and patient 6, it was observed in both cases staining in regions corresponding to glycoproteins with a high molecular weight (above 245kDa), possibly mucins. It was also noted the appearance of an intense band at molecular weight near 100kDa with HECA-452 staining in western blotting from tumor tissue of patient 5, which it was not observed with E-Ig staining, although this staining will be repeated. In the western blotting from tumor tissue of patient 6, it was observed a band below 75kDa with HECA-452 staining which it was not observable with E-Ig staining. The specific bands observed in both cases correspond to glycoproteins decorated with sLex/sLea in their surface but they are not E-SL since these bands did not appear in E-Ig staining.

Western blotting's from tumor tissues of patient 7 and patient 9 showed a similar pattern in HECA-452 staining and E-Ig staining. A band near 100kDa and a band slightly above 180kDa were observed in both cases. It was also possible to observe a staining in a region corresponding to glycoproteins with a high molecular weight, possibly mucins.

For patient 10, besides staining in high molecular weights, a band near 245kDa was detected with HECA-452 and E-Ig staining though more intense in HECA-452 staining.

Overall, staining with HECA-452 and E-Ig in patients 1, 2, 7, 9 and 10 showed a similar pattern, thus the glycoproteins that correspond to the observed bands are decorated with sLex/sLea and are at same time E-SL in these NSCLC tumor tissues (62.5% of the analyzed cases by western blotting). It is important to remember that in patient 2, there was a band slightly above 100kDa observed in E-Ig staining but not in HECA-452 staining, so this is an E-SL glycoprotein that is not decorated with sLex or sLea.

The pattern of E-SL glycoproteins between different LC samples was not identical, however 50% of analyzed patients showed a band near 100kDa and a band slightly above 180kDa.

It was also studied the presence of sLex/sLea in matched normal lysates of the previous NSCLC patients analyzed by western blotting (patient 1, 2, 3, 5, 6, 7, 9 and 10). We could not observe any specific band in any of analyzed samples with only few samples with a weak staining in the zone above of 254kDa (data not shown).

3.2.4 Expression of glycoproteins MUC1 and CD44 in tissues samples by western blotting

After analysis of sLex/sLea expression and E-SL profile in various tumor NSCLC samples, it was important to identify the proteins scaffolds of E-SL that correspond to previous detected bands by western blotting (figure 3.7). According to the observed molecular weight, two of those glycoproteins could be MUC1 (molecular weight: 200-500kDa) and CD44 (molecular weight: 80-200kDa), two glycoproteins that play an important role in cancer.

MUC1 is a transmembrane glycoprotein highly *O*-glycosylated expressed on the apical surface of most epithelial cells of many tissues including lung, stomach, pancreas, breast and several other organs. MUC1 overexpression and changes in glycosylation of this protein is often associated with carcinomas (Nath and Mukherjee, 2014; Taylor-Papadimitriou *et al.*, 1999). Furthermore, it is described as an E-SL for

other types of cancer, such as breast cancer (Nath and Mukherjee, 2014; Tarp and Clausen, 2008; Taylor-Papadimitriou *et al.*, 1999).

CD44 is the principal receptor for hyaluronic acid (an extracellular matrix glycan) and a cell-surface glycoprotein that is involved in cell adhesion, cell migration and cell-matrix interactions. Normally, it is expressed on the surface of several human cells including fibroblasts, hepatocytes, leukocytes and endothelial cells. An important glycoform of CD44 is the sialofucosylated selectin binding glycoform known as hematopoietic cell E-selectin/ L-selectin ligand (HCELL), originally found on human hematopoietic stem cells and subsequently identified on various tumor cells (Jacobs and Sackstein, 2011; Thomas *et al.*, 2008). Recent studies reported the aberrant expression of CD44 variants in various tumors including lung, breast and colon cancer.

To detect MUC1, it was used a monoclonal antibody that binds to the tandem repeat region of MUC1 of human origin. In the case of CD44, it was used an antibody anti- CD44 clone 2C5 that recognize an epitope common to all isoforms of CD44.

The same eight patients analyzed previously by western blotting were used for this analysis. Two polyacrylamide gels were made for CD44 staining. The membranes used previously for HECA-452 staining were washed and used for MUC1 staining (see Material and Methods). In each gel of CD44 staining, one lane was used as positive control normally used by the laboratory group with a sample with a known pattern of protein bands for CD44. There was no positive control for MUC1, so the potential identification of this glycoprotein was performed only based on their molecular weight. In the remaining lanes, it was placed 20µg of each tumor lysate. The membranes were stained with anti-MUC1 and anti-CD44 and after membrane revelation, it was possible to see the pattern of protein bands recognized by anti-CD44 antibody (figure 3.8). In the case of MUC1 staining, it was not observed any band or any staining in the two membranes (data not shown), probably because the quality of the antibody. Nevertheless further assays are still necessary to understand why no staining was observed.

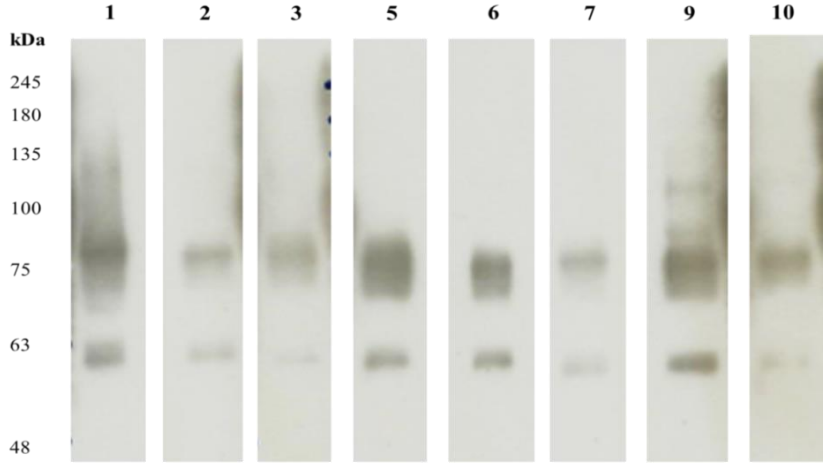


Figure 3.8: Western blotting analysis of CD44 present in tumor lysates from eight NSCLC patients. The numbers above each lane represent the number of the analyzed patient. Tumor tissue from patients was lysed and the western blotting procedure was performed as described in section "Material and Methods", using an antibody against CD44. Band sizes are indicated in kilodaltons (kDa) at the left side. Time of *films* exposition: 10min.

In all patients, the western blotting of the tumor tissue revealed a similar pattern of CD44 staining: a band near 63kDa and another band slightly above 75kDa. In the lane of patient 1 and 9, it was also possible to see the appearance of a slight band above 100kDa.

The band slightly above 75kDa corresponds possibly to hematopoietic or standard CD44 isoform (CD44s) that normally runs on SDS-PAGE gel at 80-95kDa and it is expressed on the surface of the most mammalian cell types. This protein has a theoretical molecular weight of 37kDa but the final value results from extensive glycosylation. The band near 63kDa corresponds possibly to CD44 isoform with a small glycosylation degree.

If we compare these patterns with the patterns of HECA-452 and E-SL staining, we found that they do not match, meaning CD44 probably is not decorated with sLex and/or sLea and it is also not an E-SL in these NSCLC samples.

3.2.5 Expression of carcinoembryonic antigen in tissues samples by western blotting

Another molecule very studied and used as clinical marker in some types of cancers is carcinoembryonic antigen (CEA). The main aim in this topic was to verify if this glycoprotein was one of E-SL in the NSCLC tissues.

CEA is a glycoprotein involved in cell adhesion that contains approximately 50% of carbohydrates with a molecular weight between 180-200kDa due to variations in its carbohydrate side chains. Its expression is limited in adult normal tissues but it is expressed at a high level in several tumors including colorectal carcinoma, lung adenocarcinoma and ovarian carcinoma (Beauchemin and Arabzadeh, 2013; Grunnet and Sorensen, 2012). CEA family is divided in two major groups: CEA-related molecules and pregnant-specific glycoproteins (PSG). Whereas all PSG's are secreted and restrict to placenta, CEA-related molecules contain both secreted and cell surface glycoproteins (Hammarström, 1999).

CEA is mostly detected in serum but can also be found in body fluids and tissues. Currently, it is one of the most frequent biomarkers used for colorectal cancer, serving as functional L-selectin and E-selectin ligand in this type of cancer, which may be critical to the metastatic dissemination of the tumor (Duffy, 2001; Thomas *et al.*, 2008).

To detect CEA it was used an anti-human CD66e antibody that recognizes the carcinoembryonic antigen-related cell adhesion molecule 5 (CEACAM5) also known as CEA.

The same eight patients analyzed previously by western blotting were used for this analysis. Two polyacrylamide gels were made for CEA staining. In each gel, one lane was used as positive control normally used by the laboratory group with a sample that contain CEA. In the remaining lanes, it was placed 20µg of each tumor lysate. The membranes were stained and after membrane revelation, it was possible to see the pattern of protein bands recognized by anti-CD66e antibody (figure 3.9).

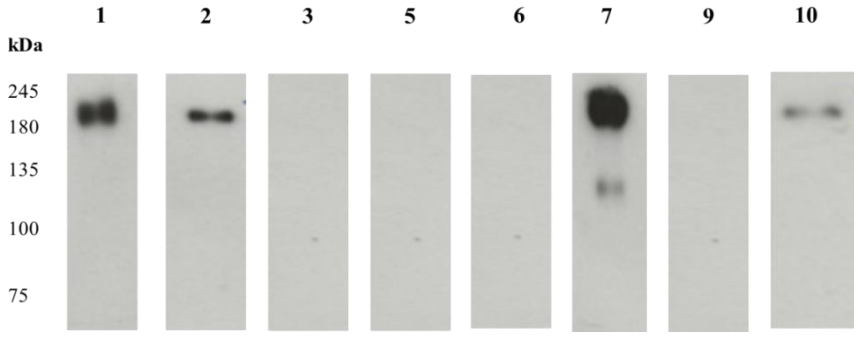


Figure 3.9: Western blotting analysis of CEA proteins present in tumor lysates from eight NSCLC patients. The numbers above each lane represent the number of the analyzed patient. Tumor tissue from patients was lysed and the western blotting procedure was performed as described in section "Material and Methods", using an antibody against CEA. Band sizes are indicated in kilodaltons (kDa) at the left side. Time of *films* exposition: 30min.

In figure 3.9, it was possible to observe the presence of CEA in 50% of NSCLC tissues more specifically in tissues from patient 1, 2, 7 and 10 with the appearance of a band with a molecular weight between 180-200kDa. Lysate from patient 7 was the one that presented a more intense staining for CEA with an intense band at 180-200kDa and also the presence of another band bellow of 135kDa that could correspond to other CEA-related molecules.

If we compared CEA staining with the pattern of HECA-452 and E-SL staining (see figure 3.10) in these patients, we found a matched band at 180-200kDa (except for patient 10) which indicate that CEA can be decorated with sLex and/or sLea antigens and be probably an E-SL in these LC tissues.

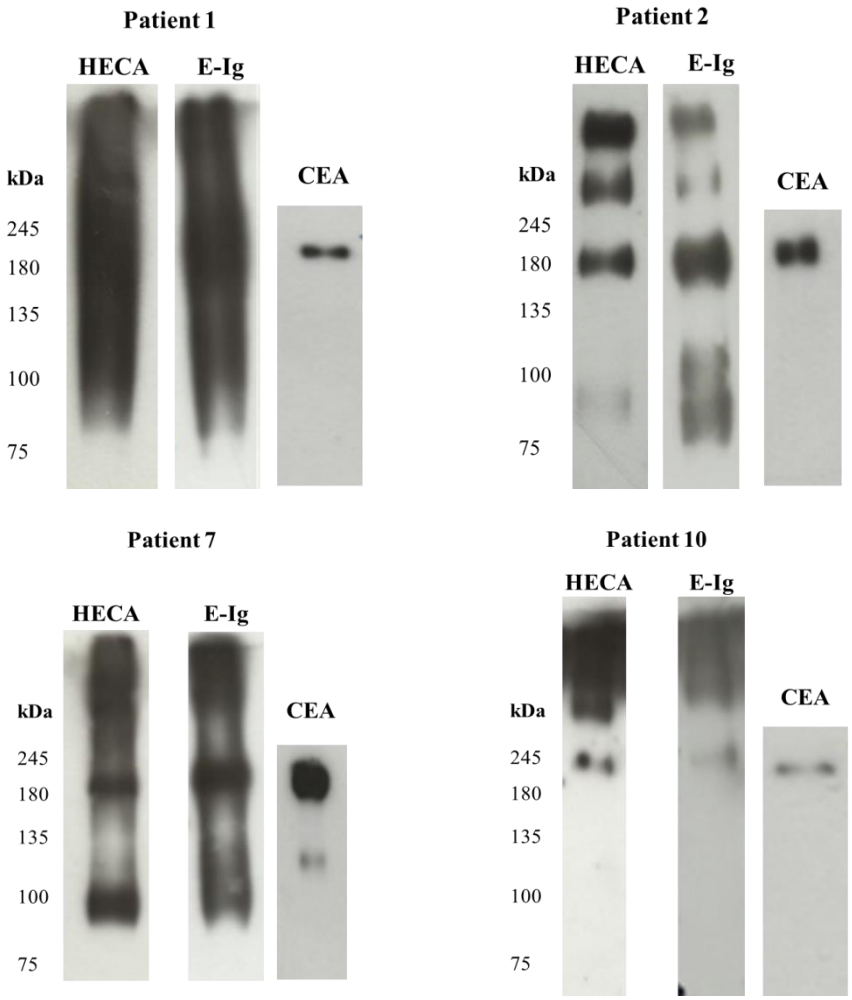


Figure 3.10: Comparison of western blotting analysis with HECA-452 staining, E-Ig staining and CEA staining in tumor lysates from four NSCLC patients. Band sizes are indicated in kilodaltons (kDa) at the left of each set.

It was also studied the presence of CEA in normal lysates of NSCLC patients analyzed by western blotting (patient 1, 2, 3, 5, 6, 7, 9 and 10). We could not observe any band in any of analyzed samples, confirming that in normal tissues the expression of CEA is reduced and limited.

As CEA molecule appeared in 50% of analyzed samples and seems to be an E-SL, it was examined if other tumor samples from NSCLC patients also expressed this glycoprotein. The chosen samples were tumor samples from patient 12, 13, 14, 16, 17, 18, and 19, the remaining patients that showed some expression of sLex/sLea antigens by dot blotting. The pattern of protein bands recognized by anti-CD66e antibody in these tumor samples is presented in figure 3.11. Only in lysates from patient 16 and 19 was possible to observe a band with a molecular weight between 180-200kDa that corresponds to CEA glycoprotein. In patient 19, the observed band had a very low intensity.

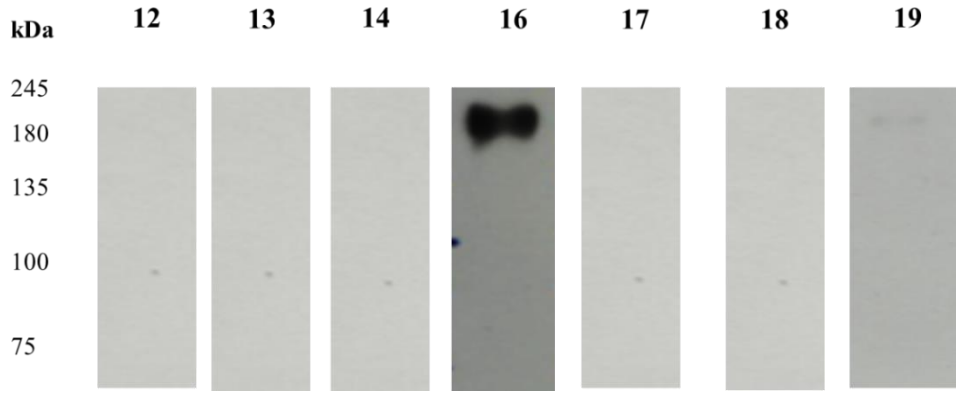


Figure 3.11: Western blotting analysis of CEA proteins present in tumor lysates from eight NSCLC patients. The numbers above each lane represent the number of the analyzed patient. Tumor tissue from patients was lysed and the western blotting procedure was performed as described in section "Material and Methods", using an antibody against CEA. Band sizes are indicated in kilodaltons (kDa) at the left side. Time of films exposition: 30min.

3.2.5.1 Confirmation of CEA as E-SL in patient 7 by immunoprecipitation of CEA and western blotting

The main aim of this experience was to confirm that CEA antigen was an E-SL in LC. For this objective, the patient number 7 was chosen because among all analyzed patients by western blotting, it was the one that presented the highest staining for CEA.

The techniques used were the immunoprecipitation of CEA from tumor lysate of patient 7, followed by western blotting with HECA-452 staining and E-Ig staining

(figure 3.12). With immunoprecipitation, CEA glycoprotein will be separate and isolate from tumor tissue lysate that contains many thousands of different proteins and after that will be analyzed by western blotting. The staining of CEA immunoprecipitate (IP-CEA) with HECA-452 and E-Ig revealed a band in a molecular weight between 180-200kDa, confirming that CEA is an E-SL decorated with sLex/sLea antigens in NSCLC tissues.

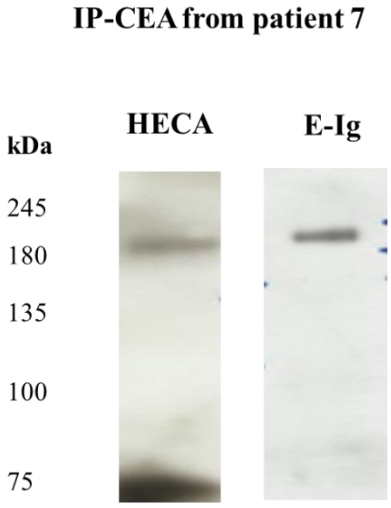


Figure 3.12: Western blotting analysis of CEA immunoprecipitate (IP-CEA) from tumor lysate of patient number 7. First lane corresponds to anti-sLex/sLea, i.e HECA-452 staining (HECA) and second lane corresponds to chimera E-Ig staining. The immoprecipitation and western blotting procedure is described in section "Material and Methods". Band sizes are indicated in kilodaltons (kDa) at the left side. Time of *films* exposition: 10min.

After confirmation of CEA as E-SL, it was studied the type of glycosylation presents on that glycoprotein. For that, it was made the immunoprecipitation of CEA from tumor lysate of patient 7 and after that the immunoprecipitate was divided: one non-treated and another one treated with an enzyme called peptide-N-glycosidase F (PNGase F). This enzyme cleaves between the innermost GlcNac and asparagine residues from *N*-glycans. PNGase F is not able to cleave *N*-glycans from glycoproteins when an α 1→3 fucose residue is linked to the innermost GlcNac residue but this modification is normally found in plants and some insects, not in animal glycoproteins.

Besides IP-CEA non-treated and treated, we also run by western blotting total tumor lysate of patient 7 non-treated and treated with the same enzyme, to understand if glycoproteins present in this tumor lysate were mostly *N*-glycans.

It was made one polyacrylamide gel for HECA-452 staining where we run IP-CEA treated and non-treated, total tumor lysate treated and non-treated and a positive control. The membranes were stained with HECA-452 and after membrane revelation, it was possible to see the pattern of protein bands recognized by antibody HECA-452 (see next figure: figure 3.13).

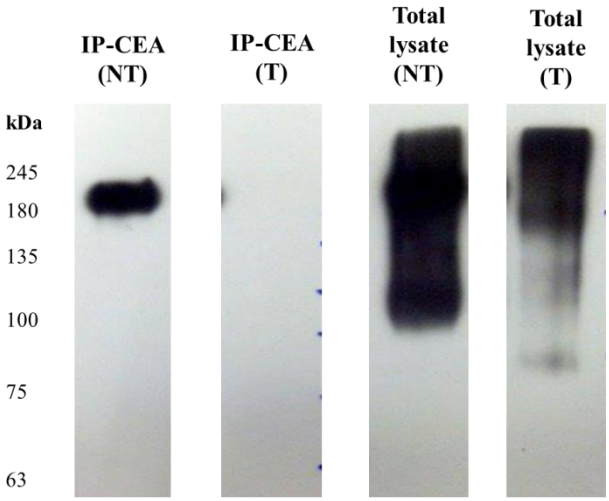


Figure 3.13: Western blotting analysis of CEA immunoprecipitate (IP-CEA) from tumor lysate of patient number 7, non-treated (NT) and treated (T) with PNGase F and also analysis of total tumor lysate of patient number 7 non-treated (NT) and treated (T) with PNGase F. Membrane was stained with antibody HECA-452 to detect sLex and sLea antigens. The immoprecipitation, treatement with the enzyme and western blotting procedure is described in section "Material and Methods". Band sizes are indicated in kilodaltons (kDa) at the left side. Time of *films* exposition: 7s.

As expected, in the lane of IP-CEA with no treatment was possible to observe the presence of a specific and intense band with a molecular weight between 180-200kDa corresponding to sLex/sLea bound to CEA antigen. With the PNGase F treatment in IP-CEA, the existing *N*-glycans in this glycoprotein were removed. Thus in second lane (IP-CEA Treated) the band that corresponds to CEA disappeared completely meaning that sLex/sLea in CEA glycoprotein is an *N*-linked glycan.

Comparing the pattern of HECA-452 staining in total lysate treated and non-treated, it was possible to observe a reduction in the intensity of the staining in total lysate treated with PNGase F, meaning that overall the sLex/sLea present in this patient are *N*-glycans.

3.3 Glycosidic characterization of lung cancer cell lines

Cancer cell lines are *in vitro* models used in laboratory to study the biology of the cancer, trying to translate these findings to clinical applications. HeLa cell line was the first cultured cancer cell line derived from cervical cancer cells in 1951.

Unlike biological samples with variability and heterogeneity, cancer cell lines represent homogeneity in tumor cells as the selective pressure of the culture conditions tends to produce a homogeneous culture of the most "strong" cancer cells. Furthermore, cell lines are able to replicate infinitely and so can offer a limitless source of material.

3.3.1 Analysis of gene expression by RT-PCR

LC cell lines are a good model to study the expression of the genes involved in biosynthesis of sLex and sLea antigens. These results will complement the results found in tissue samples of NSCLC patients. In this master thesis, three cell lines derived of NSCLC were analyzed: A549, H292 and H1299 (characteristics of these cell lines and procedure of culture and maintenance are described in section of material and methods).

The genes selected for gene expression analysis were the same genes analyzed in tissues samples of NSCLC patients: FUT3, FUT4, FUT6, FUT7, ST3GAL3, ST3GAL4, ST3GAL6, β4GALT1, GCNT1 and GALNT3. Besides these genes it was further analyzed the expression of CEA because it is a tumor biomarker and we found in our experiences with tissues samples that CEA is an important molecule that was expressed in tumor NSCLC samples and it was also an E-SL in tumor NSCLC tissues. The results of gene expression are summarized in table 3.5.

Genes that encode for FUTs presented a very low expression in all three cell lines, except for FUT3 that were expressed in a high level (relative mRNA level= 64.576) in H292 cell line.

Concerning to STs, it was observed a very low expression of these genes in all three cell lines. ST3GAL3 was the gene that showed a slightly higher expression compare with ST3GAL4 and ST3GAL6, except for H292 cell line where the level of ST3GAL3 was similar to the level of ST3GAL4.

Surprisingly, it was observed that CEA was expressed at a very low level in all three cell lines. GCNT1 was also expressed at a very low level in all three cell lines.

β4GALT1 was the gene that presented the highest level of expression in all three cell lines: H1299 (relative mRNA level = 33.855), H292 (relative mRNA level = 647.135) and A549 (relative mRNA level = 689.505).

Table 3.5: Values of gene expression of the genes analyzed by RT-PCR in three NSCLC cell lines (H292, H1299, A549). Values of gene expression were calculated using the formula $2^{-\Delta Ct}\times1000$ that infers the number of mRNA molecules of a certain gene per 1000 molecules of the average of the endogenous controls (β-actin and GAPDH).

	H292	H1299	A549
FUT3	64.576	0.104	0.058
FUT4	0.129	0.380	0.641
FUT6	0.042	0.001	0.000
FUT7	0.001	0.005	0.001
ST3GAL3	2.238	6.729	2.091
ST3GAL4	2.259	0.641	0.729
ST3GAL6	0.084	1.680	0.232
GCNT1	1.015	0.092	0.122
GALNT3	40.558	1.204	0.183
β4GALT1	647.135	33.855	689.505
CEA	0.081	0.002	3.941

3.3.2 Phenotypic analysis by flow cytometry

Analysis by flow cytometry of the three NSCLC cell lines aimed to study the expression of sLex/sLea and E-SL at the surface of cancer cells, structures as mentioned above (see section of Introduction) that have an important role in metastasis. We also want to detect the expression of some specific glycoproteins such as CEA, CD44, MUC1, MUC5AC and MUC5B which may be decorated with sialyl Lewis antigens and could be E-SL in LC. These glycoproteins are known to be expressed in LC cell lines.

The values of median intensity fluorescence (MFI) are represented in table 3.6.

Table 3.6: Values of median intensity fluorescence (MFI) obtained by flow cytometry in each NSCLC cell line for antibodies against MUC1, MUC5AC, MUC5B, CEA, CD44, sLex/sLea and E-SL. MFI values were normalized against the MFI value obtained in negative controls (see Material and Methods - section 2.4.2.4).

	MFI values		
	H292	**H1299**	**A549**
MUC1 commercial	0	0	32
MUC1 hybridoma	14	0	10
MUC5AC	0	0	24
MUC5B	6	0	59
CEA	0	0	39
CD44	11262	5.43×10^5	5.33×10^5
sLex and sLea	5706	102	24
E-selectin ligands	38639	10201	4775

It was observed that CD44 molecule was the marker that presented the highest level of expression compared with all the other markers in A549 and H1299 cell lines. The levels of CD44 expression in those cell lines were similar. In H292 cell line, although the MFI value for CD44 be positive (MFI = 11262), there was a lot of cell death in this experience, so this is not a very reliable value. In the future, this analysis will be repeat.

In the case of E-SL stained with E-Ig, H292 was the cell line that expressed more E-SL (MFI = 38639), followed by H1299 (MFI = 10201) and A549 (MFI = 4775). With a similar patter, sLex and sLea expression was in a moderate level expression in H292 cell line (MFI =5706), a low level of expression in H1299 cell line (MFI =102) and no expression in A59 cell line (MFI = 24 is a value too low to be considered expression). It was possible to observe that the cell line which demonstrated the highest expression of E-SL was also the one that demonstrated the highest expression of sLe/sLea, confirming that E-SL are mostly decorated by sLex and/or sLea antigens.

MUC1 antigen (commercial and hybridoma supernatant), MUC5AC, MUC5B and CEA showed no expression in all LC cell lines. Although some MFI values are positive, it was not considered the expression of these markers because the values were too low.

3.3.3 Functional analysis of E-selectin ligands by adhesion assay

The main aim of these experiences was to study the ability of LC cells to bind to E-selectin under dynamic/flow conditions to simulate the environment of tumor cells *in vivo*. Sometimes, though E-SL are present in cancer cells, they do not bind to E-selectin in flow conditions, because they have incomplete posttranslational modifications, which are required in order to be functional.

For each NSCLC cell line, the recombinant mouse E-selectin/CD62E Fc chimera (E-Ig) was immobilized on coated slides and after blocking to reduce non-specific interactions, LC cells were incubated above the immobilized slide under agitation. E-Ig will recognize all the E-SL, sialylated and fucosylated molecules in cancer cells that bind to the lectin domain of E-selectin.

There were two conditions in these experiences: one with cells in the presence of Ca^{2+} and another one with cells in EDTA. These two conditions are essential because the binding to E-selectin is dependent on the presence of calcium and by other way, EDTA is a chelating agent that chelates calcium molecules.

After removing the unbound cells and cell fixation, the slides were observed in a light microscope with an attached camera. A representative photograph was taken and the adhered cells were counted using ImageJ software (figure 3.14).

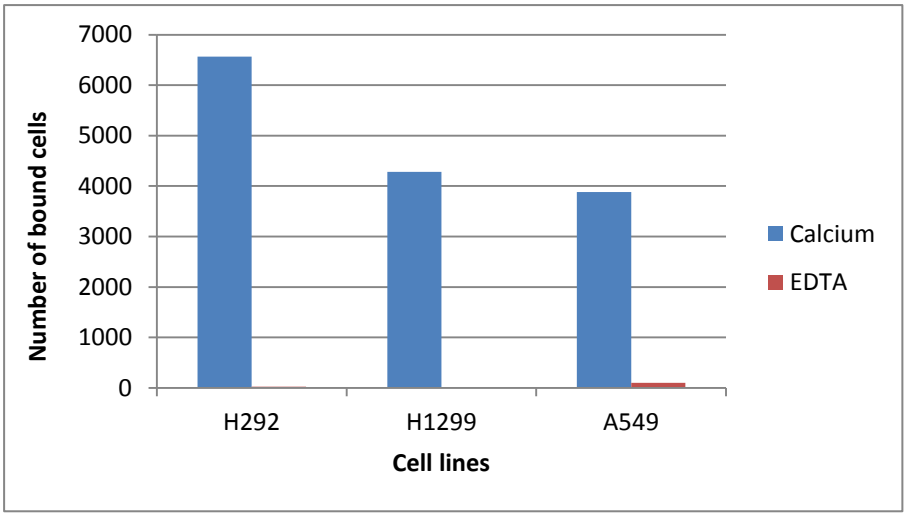


Figure 3.14: Adhesion of lung cancer cells to E-selectin. Data show values corresponding to the number of bound cells in each cell line (H292, H1299 and A549) with calcium and with EDTA. The counting of the cells was made using ImageJ 1.48v program. Parameters used in program: size of the cells (0.003 in H292 and A549 and 0.01 in H1299).

As you can see in figure 3.14, the cells that adhere more to E-Ig (binding dependent of calcium) were H292 cells (number of bound cells: 6563), followed by H1299 (number of bound cells: 4280) and A549 (number of bound cells: 3879).

In the case of cells with EDTA, it was observed a low number of bounded cells to E-Ig comparing with cells with calcium due to chelating effect of EDTA that inhibits the bonds: H292 - 24, H1299 - 9 and A549 - 99.

Together with these results it was possible to observe a positive correlation between adhesion of LC cells and the amount of E-SL and sLex/sLea that cancer cells express. Besides that, the cell line that expresses more E-SL and sLex/sLea (H292 cell line) is the one that express more FUT3, suggesting that this fucosyltransferase is probably important and essential in the biosynthesis of sLex/sLea antigens.

4. Discussion and conclusions

Glycans can impact a wide spectrum of cellular functions ranging from cell adhesion to cellular signaling (Fuster and Esko, 2005; Ohtsubo and Marth, 2006; Pinho and Reis, 2015). Consequently, alterations in glycan structures contribute significantly to changes in properties of the cells, playing an important role in cancer development and progression (Kim and Varki, 1997; Schultz, Swindall and Bellis, 2012; Stowell, Ju and Cummings, 2015). Indeed, altered glycosylation can be considered a universal feature of cancer cells.

LC is the leading cause of death related to cancer around the world. With the number of deaths increasing over time, there is an urgent need for clinical markers. Molecules including CEA, cytokeratin 19 fragment and tissue polypeptide antigens (TPA) are potential serum protein biomarkers for LC diagnosis, however they lack in sensitivity and specificity (Adamczyk, Tharmalingam and Rudd, 2012; Meany and Chan, 2011). Most candidate biomarkers focus on the polypeptide component but targeting glycans in combination with the protein scaffold will provide a greater diagnostic performance.

Sialyl Lewis antigens are one of the structures aberrantly expressed in several types of cancers, including LC (Fukuoka, Narita and Saijo, 1998; Nishida *et al.*, 1991). These antigens play an important role as E-SL since they are expressed on the surface of glycoproteins in cancer cells and can possibly contribute to adhesion of cancer cells to endothelium and subsequent metastasis (Bendas and Borsig, 2012; Witz, 2008). Nevertheless, E-SL in NSCLC is still poorly understood.

Therefore, this work had as general purpose the identification of new glycan-based biomarkers in NSCLC targeting sLex/sLea in combination with protein scaffolds and also important enzymes involved in the biosynthesis of sialyl Lewis antigens. Moreover, in this thesis it was investigated the expression of E-SL in NSCLC tissues as well as their molecular basis, in order to better understand their pathophysiological role in cancer progression.

This way, it was studied glycosidic alterations and the pattern of glycosylation in NSCLC tissues and corresponding normal tissues, as well as changes in expression of glycosidic antigens in NSCLC cell lines. Together, these results will provide useful information for the diagnostic, prognosis and immunotherapy in NSCLC.

4.1 FUT3, FUT6 and FUT7 are upregulated, while FUT4 and ST3GAL4 are downregulated in NSCLC tumor tissues compared with normal NSCLC tissues.

The results obtained in this work show that FUT3, FUT6 and FUT7 are upregulated in NSCLC tumor tissues compared with normal tissues. These results are in agreement with the fact that FUTs are overexpressed in several types of tumors, including LC (Barthel *et al.*, 2009; Dall'Olio *et al.*, 2014; Togayachi *et al.*, 1999). Since these enzymes catalyze the final step in the biosynthesis of sialyl Lewis antigens, these results suggests that the amount of sLex and sLea epitopes in patients with upregulation of these FUTs should be greater compared with patients without upregulation of these enzymes, a topic that will be addressed later.

In tumor tissues, FUT3 is expressed at high levels in opposition to FUT6 and FUT7 that are expressed at low levels. This could be explain due the fact that FUT3 is the only FUT responsible for the synthesis of sLex and sLea epitope whereas the others FUTs are only involved in sLex biosynthesis (Chen, 2011; Dall'Olio *et al.*, 2012).

We found that FUT4 is downregulated in NSCLC tumor tissues compared with normal tissues. This enzyme is known to contribute essentially to the synthesis of Le^y and Le^x and only participates in biosynthesis of sialyl Lewis antigens under especial conditions (Kannagi, 1997). Although high expression of FUT4 and FUT7 has been correlated with a poor prognosis of LC (Ogawa, Inoue and Koide, 1996) and high levels of FUT4 were observed in several types of cancer such as gastric and colorectal carcinoma (Petretti *et al.*, 1999, 2000), in our work FUT4 does not seem to be the key FUT responsible for sLex/sLea synthesis.

Besides FUT4, ST3GAL4 is also downregulated in NSCLC tumor tissues compared with normal tissues. This finding is in line with the results from other authors that observed a downregulation of this enzyme in human renal cell carcinoma and colorectal cancer tissues (Ito *et al.*, 1997; Saito *et al.*, 2002). ST3GAL4 is responsible for the sialylation of the type 2 chains terminal galactose residue. Downregulation of this gene do not seem to influence the expression of sLex and sLea, since we observed that these antigens were highly expressed in NSCLC tumor tissues as will be addressed in the 4.3 topic in this discussion. Then, other enzymes such as ST3GAL3 and ST3GAL6 are probably the responsible for the sialylation of these structures.

It is important to refer some limitations of the present study that includes the use of normal pulmonary tissue from NSCLC patients to provide "normal tissue" for comparison. Preferably, normal tissue should be derived from patients without LC in order to avoid any change of characteristics that can result from cancer tissues.

4.2 Increased expression of FUT4 is associated to a more advanced stage of NSCLC.

Despite FUT4 levels in NSCLC tumor samples are downregulated compared with NSCLC normal samples, it is observed an increase of FUT4 expression in patients with a more advanced stage of NSCLC (stage IIIA) in comparison with patients in first stage (IA and IB), which show a low expression of FUT4. Until the time, it was not reported any correlation between the stage of LC and the expression of any FUT. In fact, the only protein associated with an advanced stage of NSCLC is BIRC6 (baculoviral IAP repeat-containing 6) protein, reported in (Dong *et al.*, 2013).

Thus, FUT4 may be a biomarker for diagnosis of stages of LC since a low expression of this gene is correlated with early stages of the disease. Importantly, this possible biomarker only "works" in patients already diagnosed with LC, in which the doctors only want to know the stage of the disease, since FUT4 is expressed also at low levels in normal NSCLC tissues, as described previously in this thesis.

It was also possible to observe that patients with more of 65-age show a tendency to express high levels of all studied genes (FUT3, FUT4, FUT6, FUT7 and ST3GAL4) compared with the patients less than 65-age. However, there is a need for a larger sample size of patients to prove the significance of this result. A larger sample size would probably influence some results that approached, but failed to reach, statistical significance. In this specific case, with more patients, we could probably infer a relationship between the age of the NSCLC patients and the expression of the genes involved in the biosynthesis of sLex and sLea antigens.

4.3 sLex/sLea are more expressed in NSCLC tumor tissues than normal tissues.

The majority of NSCLC tumor tissues expressed more sLex and sLea in comparison with NSCLC normal tissues as previously shown in serum from NSCLC patients (Satoh *et al.*, 1998, 2002) and also NSCLC biopsies (Fukuoka *et al.*, 1998).

It is known that E-selectins expressed on activated endothelial cells exhibit a preference for ligands with sLex/sLea motif in the surface of cancer cells. Previous studies show that overexpression of sialyl Lewis antigens on cancer cells is associated with a poor prognosis due to enhanced metastatic phenotype as well as with increased ability to adhere to E-selectin (Barthel *et al.*, 2007; Häuselmann and Borsig, 2014). Thereby, our results suggest that LC patients that strongly express sLex/sLea in their tumor cell surface would have cancer cells with higher ability to bind to E-selectins, facilitating tumor proliferation and promoting hematogenous metastasis of LC. Additionally, overexpression of both antigens will alter the immune homeostasis of the mucous membrane, favoring cancer progression (Kannagi *et al.*, 2010).

Furthermore, it is well known that sLex is preferentially expressed on lung, ovary, liver, kidney and breast cancer, being expected that its expression would contribute more to LC progression than sLea (Kannagi, 1997). However, more studies are needed to clarify this matter.

4.4 Increased expression of ST3GAL3 and GCNT1 is related with enhanced expression of sLex/sLea antigens in tumor NSCLC tissues.

One of the main mechanisms driving sLex and sLea overexpression in tumor cells is the increased expression of glycosyltransferases, particularly STs and FUTs that are involved in the biosynthesis of these antigens. Normally, the overexpression of a carbohydrate structure is multifactorial. Some of these alternative mechanisms for altered expression of glycans include competition between glycosyltransferases, alteration of glycosidases expression, altered expression of sugar and sugar nucleotide transporters and also masking of sugar structures by substituent groups (Dall'Olio *et al.*, 2012; Pinho and Reis, 2015; Stowell *et al.*, 2015).

In this thesis, increased expression of ST3GAL3 and GCNT1 in tumor NSCLC tissues correlates with higher sLex and sLea levels. Surprisingly, these genes are not

upregulated in NSCLC and the overexpression of sLex/sLea on tumor NSCLC tissues is not correlated with any single glycosyltransferase gene that is upregulated in tumor NSCLC (FUT3, FUT6 and FUT7), as mentioned above. A likely explanation for ST3GAL3 is based on competition between FUTs and enzymes synthesizing alternative structures, such as Sda antigen or sialyl 6-sulfo Lewis^{x} antigen as described for human colorectal cancer and gastrointestinal cancer (Izawa *et al.*, 2000; Kawamura *et al.*, 2005; Malagolini *et al.*, 2007). However, sLex/sLea overexpression in LC remains unclear.

Concerning to ST3GAL3, it is important to remember that within all studied STs, ST3GAL3 is the one that presents the highest level of expression, being the most likely ST that will directly influence the expression of these antigens in LC. ST3GAL3 is responsible for the transference of sialic acid residue especially to Galβ 1→3 GlcNAc and Galβ 1→4 GlcNAc on glycoproteins and glycolipids, suggesting that this enzyme is involved in the generation of sLex and sLea determinants. A previous study shows that the overexpression of ST3GAL3 in pancreatic cancer is correlated with an increase in sLex antigen and an enhancement of their potential to metastasize in animal models (Banerjee and Hart, 2014). In LC, this association has not been reported.

GCNT1 catalyzes the transference of GlcNAc residue to form the β1→6 linkage at GalNac of core-1 *O*-linked glycans. Although it is an enzyme involved in early steps of *O*-glycosylation, these results demonstrate that the expression of GCNT1 is correlated with expression of sLex and sLea antigens. The expression of this gene has been associated with malignant potential of several types of cancer including colorectal cancer, testicular germ cell tumor and also pulmonary carcinoma (Hatakeyama *et al.*, 2010; Machida *et al.*, 2001).

4.5 E-selectin ligands in lung cancer: CEA is an *N*-glycan E-SL in NSCLC while CD44 and MUC1 don’t seem to be.

E-selectin is expressed on endothelial cells and initially its physiological role was only associated with the recruitment of leukocytes to stimulated endothelium during inflammation. Nowadays, it is recognized as an important lectin that interacts with cell-surface glycoconjugates in tumor cells denominated E-SL mediating tethering, rolling and adhesion of tumor cells to endothelium. This adhesion can possibly promote the formation of metastasis in cancer. E-selectin is not constitutively expressed on

endothelial cells: its expression is strongly induced by inflammatory cytokines such as IL-1β, IL-1α or TNF-α (Kannagi *et al.*, 2004).

So far, several glycoproteins scaffolds of E-SL have been identified in several types of cancer, being the most prominent the CD44 glycoform, HCELL, as E- and L-selectin ligand on colon cancer and CEA on colon and prostate cancer (Burdick *et al.*, 2012; Thomas *et al.*, 2008). Until now, it was not described any glycoprotein as an E-SL in NSCLC. However, in small cell lung cancer have already been identify E- and P-selectin ligands, one of them CEA (Heidemann *et al.*, 2014; Li *et al.*, 2001).

In this master thesis, CEA was identified as an E-SL in 50% of NSCLC patients that expressed this molecule. Besides that, it was demonstrated that CEA is mainly decorated with sialyl Lewis antigens, sLex and/or sLea. Despite of tetrasacharides sLex and sLea being the main recognition motifs for E-selectin, there are cancer cells that did not express these molecules but still adhere strongly to E-selectin. There are other structures than can bind to E-selectin but in a smaller extent such as VIM-2 epitope, internally fucosylated α(2,3)-sialylated glycolipids and structures related to sLex/sLea antigens in which the sialic acid is substituted by a sulfate group (Kannagi, 1997; Vestweber and Blanks, 1999).

Additionally, we demonstrated that CEA in NSCLC patients is constituted by *N*-glycans. Some studies suggest an important role of *N*-glycans in binding to E-selectin and others suggest that binding to E-selectin is largely dependent of *O*-glycans (Mondal, Buffone and Neelamegham, 2013). Despite only one patient has been studied regarding the type of glycosylation present in CEA, we extrapolate that all the other NSCLC patients who expressed CEA molecule also have the same type of glycosylation.

As previously shown, it was not possible to detect CEA molecule in all patients analyzed. This can be explained by the variability of results that is probably originated from normal biological variability of the different patients with particular intrinsic characteristics.

CEA is a glycoprotein with a limited expression in normal tissues but detected in high levels in tumors with epithelial origin including lung adenocarcinoma, gastric carcinoma and colorectal cancer. It is currently used as tumor biomarker in colorectal tumors (Duffy, 2001). In NSCLC, CEA expression has been mainly studied in serum with its levels been reported to be correlated with poor therapeutic response and

survival, advanced stage of the disease and also early relapse (Grunnet and Sorensen, 2012; Wang *et al.*, 2012). Overall, our data suggests a new function for this molecule in NSCLC tissues: CEA is an E-SL in NSCLC that may facilitate adhesion to endothelium and consequently cancer metastasis. Further studies with larger cohorts of patients are required to confirm his role as an E-SL, to investigate CEA as potential therapeutic target to combat metastasis in LC and also to confirm their usefulness as a diagnostic marker in NSCLC patients.

Besides CEA, it was also studied CD44 molecule and MUC1 as potential E-SL in NSCLC patients. However, it was not possible to find this association in our results.

4.6 Enhanced expression of sLex/sLea on lung cancer cells which may be attributed to elevated levels of FUT3 correlates with increased ability to adhere to E-selectin.

Our results based on the study of LC cell lines suggest that the elevated expression of E-SL and also sialyl Lewis antigens (sLex and sLea) found by flow cytometry is correlated with high levels of FUT3 expression found by RT-PCR. Consequently, these facts are also correlated with increased ability to adhere to E-selectin, as observed in adhesion assays.

In H292 cell line, all FUTs genes analyzed by RT-PCR show a low expression level with exception of FUT3. It was also possible to observe high levels of expression of β4GALT1 and GALNT3 but these enzymes are found in early steps of sialyl Lewis antigens biosynthetic pathway, having probably a less prominent role. As previously shown in our results, FUT3 is overexpressed in tumor NSCLC tissues when compared with normal tissues. Taken together, FUT3 seem to be the main enzyme responsible for the synthesis of sLex/sLea antigens.

Surprisingly, it was not possible to identify protein scaffolds of E-SL (MUC1, MUC5AC, MUC5B and CEA) by flow cytometry although they have been reported to be expressed in LC cell lines and NSCLC patients (Bührens *et al.*, 2009; Croce *et al.*, 1999; Matsubara *et al.*, 2000), with exception of CD44 molecule that is expressed at high levels in the three cell lines studied. Curiously, CEA molecule was not detected in any of NSCLC cell lines by genetic expression or either by phenotypic analysis, although it has been demonstrated in this thesis its important role in NSCLC tumor

tissues. One possible explanation for this is related to the fact that cell lines may have undergone many alterations with its establishment and culture, mainly at a glycosylation level, not being the most appropriate and reliable model in this study. Thus the importance of using patient's samples which is a much more physiological model, closer to the characteristics of the disease.

Regarding the adhesion assays in dynamic conditions, H292 show greater adhesion to E-selectins compared to all other cell lines that probably results from the higher expression of E-SL and sLex/sLea antigens as previously showed in flow cytometry results. Importantly, levels of sLex/sLea expression do not follow the levels of expression of E-SL confirming that not all E-SL are decorated by sLex / sLea structures, as referred previously.

In the other cell lines (H1299 and A549), the tendency is the same: increased expression of E-SL and sLex/sLea means a higher adhesion to E-selectin. In gene expression analysis of H1299 and A549, all the genes show low levels of expression except β4GALT1. Although in these cases is not possible to find a relationship with FUT3, we found a relation between E-SL and sLex/sLea expression with adhesion to selectins. H1299 cell line presents higher levels of E-SL and sLex/sLea compared to A549, thus H1299 cells can adhere more to E-selectins than A549 cells.

In previous studies it was observed that cancer cells with high E-selectin binding activity have improved ability to adhere to endothelium and extravagate into tissues, thus showing higher metastatic potency and poor prognosis of cancer (Häuselmann e Borsig, 2014).

Together, these results suggest that H292 cell line is the best model to study sialyl Lewis antigens and E-SL in NSCLC. Additionally, it suggests that FUT3 is an important target to modulate E-SL expression.

4.7 Future perspectives

During this work, we started adhesion assays of tumor and normal tissues lysates of NSCLC patients to E-selectin, in order to demonstrate the functionality of E-selectin ligands in LC patients, under flow conditions. These assays are performed using two cell lines (CHO-E and CHO-Mock), derived from CHO (Chinese Hamster Ovary) cells. CHO- E cells overexpress E-selectins and are derived from the stably transfection of

CHO cells with cDNA encoding full-length E-selectin. By other way, CHO-mock cells are derived from CHO cells transfected only with an empty vector. These experiences are still in optimization but preliminary results show us that overexpression of E-SL in tumor patients enables a higher binding to E-selectins. Therefore, this is one of the tests to continue in the future.

Since it was demonstrated the presence of CEA in LC samples as E-SL, it would be interesting to study the expression of this marker by RT-PCR on tumor and normal samples from NSCLC patients in whom there is still tissue to analyze.

Another of the future aims is to identify new E-SL in LC, studying further the role in tumor cell migration and metastasis. MUC5AC, MUC5B and SEZ6L2 (seizure related 6 homolog (mouse) - like 2 (alias PSK-1)) are three possible protein candidates (Croce *et al.*, 1999; Ishikawa *et al.*, 2006). In addition, the identification of E-SL in NSCLC cell lines by western blotting is ongoing.

For the cell lines, the next aim is the manipulation of LC cell lines with inhibitors of glycosidic structures altered in NSCLC. First of all, we propose to inhibit the expression of all FUTs with a fluorinated fucose (2-fluorofucose). In our group, tests using this inhibitor abrogating the expression of E-SL in breast cancer cells, demonstrated that fucosylation inhibition decreases substantially breast cancer cell binding to E-selectin and subsequent migration. This inhibitor is not specific for any FUT, so we want also to use RNAi approach (e.g. siRNA and antisense) to block FUT3 (FUT overexpressed in NSCLC patients and expressed at high levels in H292 cell line) gene function through sequence-specific post-transcriptional gene silencing and in the down-regulation of gene expression. We will evaluate the silencing efficacy of the genes and also evaluate the functional effect of these inhibitions in LC cell proliferation and transmigration through endothelium layer, using appropriated *in vitro* and *in vivo* assays developed by our group.

It was also initiated in collaboration with Professor Fabio Dall'Olio from University of Bologna (Italy) the analysis of FUTs activities involved in the biosynthesis of sialyl Lewis antigens in normal and tumor samples from NSCLC patients. It would also be very important to obtain information about the survival time of analyzed patients in order to understand if there is any correlation with some of the genes / markers studied.

4.8 General conclusions

This work demonstrated the great importance of E-SL in LC. Overall, NSCLC tissues show an overexpression of sLex/sLea antigens that correlates with expression of ST3GAL3 and GCNT1. It was also observed an overexpression of key FUTs (FUT3, FUT6 and FUT7) in NSCLC tumor tissues that surprisingly seem to not be correlated with sLex/sLea expression. For the first time, it was described a relation between FUT4 levels and NSCLC TNM stage. Increased expression of FUT4 is associated with a more advanced stage of NSCLC patients, suggesting that FUT4 can be a future biomarker for diagnosis of stages in LC, although there is a need for further studies.

By western blotting, we identified the CEA molecule as an E-SL decorated with sLex and/or sLea in NSCLC patients that could facilitate the adhesion of NSCLC cells to vascular endothelium, enhancing their metastatic potency. We tried to identify other E-SL proteins such as CD44 and MUC1 but that was not possible.

Results from NSCLC cell lines contributed to understand that a higher expression of E-SL and sLex/sLea in LC can be attributed to elevated levels of FUT3. This is also correlated with increased adherence to E-selectins, suggesting that E-SL are important for cancer cell invasion. These data suggest that FUT3 is an important target to modulate E-SL expression.

In this work, it was also performed the optimization of some techniques such as RNA extraction from LC tissues and adhesion assays of NSCLC lysates to cells that express E-selectin under flow conditions. This last technique is still in optimization.

In summary, this thesis contributed to a better understanding of the glycosidic changes and molecules that can influence tumor progression of LC, allowing identifying in the future new diagnosis/prognosis biomarkers and potential therapeutic targets for NSCLC.

5. References

ADAMCZYK, B., THARMALINGAM, T., RUDD, P. M. - Glycans as cancer biomarkers. Biochimica et Biophysica Acta - General Subjects. ISSN 03044165. 1820, 9 (2012), 1347–1353.

AEBI, M. - N-linked protein glycosylation in the ER. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research. ISSN 01674889. 1833, 11 (2013), 2430–2437.

AISNER, D. L., MARSHALL, C. B. - Molecular pathology of non-small cell lung cancer: A practical guide. American Journal of Clinical Pathology. ISSN 00029173. 138, 3 (2012), 332–346.

ALBERG, A. J., BROCK, M. V., SAMET, J. M. - Epidemiology of lung cancer: Looking to the future. Journal of Clinical Oncology. ISSN 0732183X. 23, 14 (2005), 3175–3185.

American Cancer Society - Cancer Facts & Figures 2014 [online], [Consult. August 25. 2015]. Available from WWW:<http://www.cancer.org/acs/groups/content/@research/documents/webcontent/acspc-042151.pdf.

BANERJEE, P., HART, G. W. - Handbook of Glycosyltransferases and Related Genes [Online]. 2nd ed. Japan: SpringerReference, 2014 [Consult. September 3 2015]. Available from WWW:<http://link.springer.com/book/10.1007%2F978-4-431-54240-7>. ISBN 978-4-431-54239-1.

BARTHEL, S. R. [et al.] - Targeting selectins and selectin ligands in inflammation and cancer. Expert opinion on therapeutic targets. ISSN 1472-8222. 11, 11 (2007), 1473–1491.

BARTHEL, S. R. [et al.] - Alpha 1,3 fucosyltransferases are master regulators of prostate cancer cell trafficking. Proceedings of the National Academy of Sciences of the United States of America. ISSN 0027-8424. 106, 46 (2009), 19491–19496.

BEAUCHEMIN, N., ARABZADEH, A. - Carcinoembryonic antigen-related cell adhesion molecules (CEACAMs) in cancer progression and metastasis. Cancer and Metastasis Reviews. ISSN 01677659. 32, 3-4 (2013), 643–671.

BENDAS, G., BORSIG, L. - Cancer cell adhesion and metastasis: Selectins, integrins, and the inhibitory potential of heparins. International Journal of Cell Biology. ISSN 16878876. (2012).

BROCKHAUSEN, I. - Pathways of O-glycan biosynthesis in cancer cells. Biochimica et Biophysica Acta (BBA)-General Subjects. ISSN 03044165. 1473, 1 (1999), 67–95.

BÜHRENS, R. I. [et al.] - Protein expression in human non-small cell lung cancer: a systematic database. Pathobiology : journal of immunopathology, molecular and cellular biology. ISSN 1423-0291. 76, 6 (2009), 277–285.

BURDICK, M. M. [et al.] - Expression of E-selectin ligands on circulating tumor cells: cross-regulation with cancer stem cell regulatory pathways? Frontiers in oncology. ISSN 2234-943X. 2, 103 (2012).

CARVALHO, A. S. [et al.] - Differential expression of α-2, 3-sialyltransferases and α-1, 3/4-fucosyltransferases regulates the levels of sialyl Lewis a and sialyl Lewis x in gastrointestinal carcinoma cells. The international journal of biochemistry & cell biology. ISSN 1878-5875. 42, 1 (2010), 80–89.

CERAMI, E. [et al.] - The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer discovery. ISSN 21598274. 2, 5 (2012), 401–404.

CHEN, H.-L. - The Molecular Immunology of Complex Carbohydrates-3 [Online]. 1st ed. US: Springer, 2011, [Consult. August 24 2015]. Chapter 4 -Lewis glyco-epitopes: structure, biosynthesis, and functions. Availabre from WWW:< http://link.springer.com/chapter/10.1007%2F978-1-4419-7877-6_4>. ISBN 978-1-4419-7877-6.

CROCE, M. V. [et al.] - Identification and characterization of different subpopulations in a human lung adenocarcinoma cell line (A549). Pathology oncology research : POR. ISSN 1219-4956. 5, 3 (1999), 197–204.

CRUZ, C. S. D., TANOUE, L. T., MATTHAY, R. A. - Lung cancer: epidemiology, etiology, and prevention. Clinics in chest medicine. ISNN 02725231 32, 4 (2011) 605–644.

DALL'OLIO, F. - Protein glycosylation in cancer biology: an overview. Clinical molecular pathology. ISSN 1472-4146. 49, 3 (1996), M126.

DALL'OLIO, F. [et al.] - Mechanisms of cancer-associated glycosylation changes. Front Biosci. ISSN 1093-9946. 17, 1 (2012), 670–699.

DALL'OLIO, F. [et al.] - Sialosignaling: sialyltransferases as engines of self-fueling loops in cancer progression. Biochimica et Biophysica Acta (BBA)-General Subjects. ISSN 18728006. 1840, 9 (2014), 2752–2764.

DENNIS, J. W., GRANOVSKY, M., WARREN, C. E. - Glycoprotein glycosylation and cancer progression. Biochimica et Biophysica Acta (BBA)-General Subjects. ISSN 03044165. 1473, 1 (1999), 21–34.

DOEBELE, R. C. [et al.] - Oncogene status predicts patterns of metastatic spread in treatment-naive nonsmall cell lung cancer. Cancer. ISSN 0008543X. 118, 18 (2012), 4502–4511.

DONG, X. [et al.] - Elevated expression of BIRC6 protein in non-small-cell lung cancers is associated with cancer recurrence and chemoresistance. Journal of Thoracic Oncology. ISSN 1556-1380. 8, 2 (2013), 161–70.

DOSAKA-AKITA, H. [et al.] - N-acetylgalactosaminyl transferase-3 is a potential new marker for non-small cell lung cancers. British journal of cancer. ISSN 0007-0920. 87, 7 (2002), 751–755.

DUFFY, M. J. - Carcinoembryonic antigen as a marker for colorectal cancer: Is it clinically useful? Clinical Chemistry. ISSN 00099147. 47, 4 (2001), 624–630.

FREEZE, H. H. - Genetic defects in the human glycome. Nature Reviews Genetics. ISSN 1471-0056. 7, 7 (2006), 537–551.

FUKUOKA, K., NARITA, N., SAIJO, N. - Increased expression of sialyl Lewis x antigen is associated with distant metastasis in lung cancer patients: Immunohistochemical study on bronchofiberscopic biopsy specimens. Lung Cancer. ISSN 01695002. 20, 2 (1998), 109–116.

FUSTER, M. M., ESKO, J. D. - The sweet and sour of cancer: glycans as novel therapeutic targets. Nature reviews Cancer. ISSN 1474-175X. 5, 7 (2005), 526–542.

GAO, J. [et al.] - Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Science signaling. ISSN 1937-9145. 6, 269 (2013), pl1–pl1.

GHAZARIAN, H., IDONI, B., OPPENHEIMER, S. B. - A glycobiology review: carbohydrates, lectins and implications in cancer therapeutics. Acta histochemica. ISSN 00651281. 113, 3 (2011), 236–247.

GOUT, S., TREMBLAY, P-L.,HUOT, J. - Selectins and selectin ligands in extravasation of cancer cells and organ selectivity of metastasis. Clinical & experimental metastasis. ISSN 02620898. 25, 4 (2008), 335–344.

GROOT, P. DE, MUNDEN, R. F. - Lung cancer epidemiology, risk factors, and prevention. Radiol Clin N Am. ISSN 1557-8275. 50, 5 (2012), 863–76.

GRUNNET, M., SORENSEN, J. B. - Carcinoembryonic antigen (CEA) as tumor marker in lung cancer. Lung cancer (Amsterdam, Netherlands). ISSN 1872-8332. 76, 2 (2012), 138–43.

HADLEY, B. [et al.] - Structure and function of nucleotide sugar transporters: Current progress. Computational and Structural Biotechnology Journal. ISSN 20010370. 10, 16 (2014), 23–32.

HAMMARSTRÖM, S. - The carcinoembryonic antigen (CEA) family: structures, suggested functions and expression in normal and malignant tissues. Seminars in cancer biology. ISSN 1044-579X. 9, 2 (1999), 67–81.

HATAKEYAMA, S. [et al.] - Core 2 N-acetylglucosaminyltransferase-1 expression induces aggressive potential of testicular germ cell tumor. International Journal of Cancer. ISSN 00207136. 127:5 (2010), 1052–1059.

HÄUSELMANN, I., BORSIG, L. - Altered tumor-cell glycosylation promotes metastasis. Frontiers in oncology. ISSN 2234-943X. 4, February (2014), 28.

HEIDEMANN, F. [et al.] - Selectins mediate small cell lung cancer systemic metastasis. PLoS ONE. ISSN 19326203. 9, 4 (2014).

ISHIKAWA, N. [et al.] - Characterization of SEZ6L2 cell-surface protein as a novel prognostic marker for lung cancer. Cancer Science. ISSN 13479032. 97, 8 (2006), 737–745.

ITO, H. [et al.] - Altered mRNA expression of specific molecular species of fucosyl- and sialyl-transferases in human colorectal cancer tissues. International Journal of Cancer. ISSN 00207136. 71, 4 (1997), 556–564.

IZAWA, M. [et al.] - Expression of sialyl 6-sulfo Lewis X is inversely correlated with conventional sialyl Lewis X expression in human colorectal cancer. Cancer Research. ISSN 00085472. 60, 5 (2000), 1410–1416.

JACOBS, P. P., SACKSTEIN, R. - CD44 and HCELL: Preventing hematogenous metastasis at step 1. FEBS Letters. ISSN 00145793. 585, 20 (2011), 3148–3158.

KANNAGI, R. - Carbohydrate-mediated cell adhesion involved in hematogenous metastasis of cancer. Glycoconjugate Journal. ISSN 02820080. 14, 5 (1997), 577–584.

KANNAGI, R. [et al.] - Carbohydrate-mediated cell adhesion in cancer metastasis and angiogenesis. Cancer Science. ISSN 13479032. 95, 5 (2004), 377–384.

KANNAGI, R. [et al.] - Altered expression of glycan genes in cancers induced by epigenetic silencing and tumor hypoxia: Clues in the ongoing search for new tumor markers. Cancer Science. ISSN 13479032. 101, 3 (2010), 586–593.

KAWAMURA, Y. I. [et al.] - Introduction of Sda carbohydrate antigen in gastrointestinal cancer cells eliminates selectin ligands and inhibits metastasis. Cancer Research. ISSN 00085472. 65, 14 (2005), 6220–6227.

KIM, Y. J., VARKI, A. - Perspectives on the significance of altered glycosylation of glycoproteins in cancer. Glycoconjugate Journal. ISSN 02820080. 14, 5 (1997) 569–576.

KITAMURA, H. [et al.] - Molecular and Genetic Pathogenesis of Lung Cancer: Differences Between Small-Cell and Non-Small-Cell Carcinomas. The Open Pathology Journal. ISSN 18743757. 2, 1 (2008), 106–114..

KOBATA, A., AMANO, J. - Altered glycosylation of proteins produced by malignant cells, and application for the diagnosis and immunotherapy of tumours. Immunology and cell biology. ISSN 08189641. 83, 4 (2005), 429–439.

LANGER, C. J. - Emerging role of epidermal growth factor receptor inhibition in therapy for advanced malignancy: Focus on NSCLC. International Journal of Radiation Oncology Biology Physics. ISSN 03603016. 58, 3 (2004), 991-1002

LÄUBLI, H., BORSIG, L. - Selectins promote tumor metastasis. Seminars in cancer biology. ISSN 1044579X. 20, 3 (2010), 169-177.

LEY, K. [et al.] - Getting to the site of inflammation: the leukocyte adhesion cascade updated. Nature Reviews Immunology. ISSN 1474-1733. 7, 9 (2007), 678–689.

LI, L. [et al.] - Characterization of glycoprotein ligands for P-selectin on a human small cell lung cancer cell line NCI-H345. Biochem Biophys Res Commun. ISSN 0006291X. 288, 3 (2001), 637–644.

LIVAK, K. J., SCHMITTGEN, T. D. - Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods (San Diego, Calif.). ISSN 1046-2023. 25, 4 (2001), 402–408.

MACHIDA, E. [et al.] - Clinicopathological significance of core 2 β1,6-N-acetylglucosaminyltransferase messenger RNA expressed in the pulmonary adenocarcinoma determined by in situ hybridization. Cancer Research. ISSN 00085472. 61, 5 (2001), 2226–2231.

MALAGOLINI, N. [et al.] - Biosynthesis and expression of the Sda and sialyl Lewis x antigens in normal and cancer colon. Glycobiology. ISSN 09596658. 17, 7 (2007), 688–697.

MATSUBARA, Y. [et al.] - Expression of CD44 variants in lung cancer and its relationship to hyaluronan binding. The Journal of international medical research. ISSN 0300-0605. 28, 2 (2000), 78–90.

MEANY, D. L., CHAN, D. W. - Aberrant glycosylation associated with enzymes as cancer biomarkers. Clinical proteomics. ISSN 1559-0275. 8, 1 (2011), 7.

MIRSADRAEE, S. - The 7th lung cancer TNM classification and staging system: Review of the changes and implications. World Journal of Radiology. ISSN 1949-8470. 4, 4 (2012), 0.

MIZUGUCHI, S. [et al.] - Clinical value of serum cytokeratin 19 fragment and sialyl-Lewis x in non-small cell lung cancer. The Annals of thoracic surgery. ISSN 00034975. 83, 1 (2007), 216–221.

MOLLICONE, R., CAILLEAU, A., ORIOL, R. - Molecular genetics of H, Se, Lewis and other fucosyltransferase genes. Transfusion clinique et biologique. ISSN 12467820. 2, 4 (1995), 235–242.

MONDAL, N., BUFFONE, A., NEELAMEGHAM, S. - Distinct glycosyltransferases synthetize E-selectin ligands in human versus mouse leukocytes. Cell Adhesion & Migration. ISSN 1933-6926. June (2013), 1–5.

MOREMEN, K. W., TIEMEYER, M., NAIRN, A. V - Vertebrate protein glycosylation: diversity, synthesis and function. Nature Reviews Molecular Cell Biology. ISSN 1471-0072. 13, 7 (2012), 448–462.

NAKAMORI, S. [et al.] - Increased expression of sialyl Lewisx antigen correlates with poor survival in patients with colorectal carcinoma: clinicopathological and immunohistochemical study. Cancer research. ISSN 1538-7445. 53, 15 (1993), 3632–3637.

NATH, S., MUKHERJEE, P. - MUC1: A multifaceted oncoprotein with a key role in cancer progression. Trends in Molecular Medicine. ISSN 1471499X. 20, 6 (2014), 332–342.

NISHIDA, K. [et al.] - Elevated tissue concentrations of sialyl Lex-i in cancerous tissues compared with those in noncancerous tissues of various organs. Cancer. ISSN 0008543X. 68, 1 (1991), 111–117.

OGAWA, J., INOUE, H., KOIDE, S. - Expression of alpha-1,3-fucosyltransferase type IV and VII genes is related to poor prognosis in lung cancer. Cancer research. ISSN 0008-5472. 56, 2 (1996), 325–329.

OHTSUBO, K., MARTH, J. D. - Glycosylation in cellular mechanisms of health and disease. Cell. ISSN 00928674. 126, 5 (2006), 855–867.

PALCIC, M. M. - Methods in Enzymology [Online]. [S.l.]:Academic Press, Inc.,1994, [Consult. August 24 2015]. Chapter 18 - Glycosyltransferases in glycobiology. Available from WWW:< http://www.sciencedirect.com/science/article/pii/0076687994300208>. ISBN: 978-0-12-802293-1.

PÉREZ-GARAY, M. [et al.] - α2, 3-Sialyltransferase ST3Gal IV promotes migration and metastasis in pancreatic adenocarcinoma cells and tends to be highly expressed in pancreatic adenocarcinoma tissues. The international journal of biochemistry & cell biology. ISSN 13572725. 45, 8 (2013), 1748–1757.

PETRETTI, T. [et al.] - Altered mRNA expression of glycosyltransferases in human gastric carcinomas. Biochimica et Biophysica Acta - General Subjects. ISSN 03044165. 1428, 2-3 (1999), 209–218.

PETRETTI, T. [et al.] - Altered mRNA expression of glycosyltransferases in human colorectal carcinomas and liver metastases. Gut. ISSN 00175749. 46, 3 (2000), 359–366.

PINHO, S. S., REIS, C. A. - Glycosylation in cancer: mechanisms and clinical implications. Nature Reviews Cancer. ISSN 1474-175X. 15, 9 (2015), 540–555.

RADCLIFF, G., JAROSZESKI, M. J. - Basics of flow cytometry. Methods in molecular biology (Clifton, N.J.). ISSN 1064-3745. 91, (1998), 1–24.

RAMAN, R. [et al.] - Glycomics: an integrated systems approach to structure-function relationships of glycans. Nature Methods. ISSN 1548-7091. 2, 11 (2005), 817–824.

RECK, M. [et al.] - Management of non-small-cell lung cancer: recent developments. The Lancet. ISSN 1474-547X. 382, 9893 (2013), 709–719.

REYMOND, N., D'ÁGUA, B. B., RIDLEY, A. J. - Crossing the endothelial barrier during metastasis. Nature Reviews Cancer. ISSN 1474-1768. 13, 12 (2013), 858–870.

RIIHIMÄKI, M. [et al.] - Metastatic sites and survival in lung cancer. Lung Cancer. ISSN 01695002. 86, 1 (2014), 78–84.

SAITO, S. [et al.] - Clinical significance of ST3Gal IV expression in human renal cell carcinoma. Oncology reports. ISSN 1021-335X. 9, 6 (2002), 1251–5.

SATOH, H. [et al.] - Elevated serum sialyl Lewis X-i antigen levels in non-small cell lung cancer with lung metastasis. Respiration. ISSN 00257931. 65, 4 (1998), 295–298.

SATOH, H. [et al.] - Serum sialyl Lewis X-i antigen in lung adenocarcinoma and idiopathic pulmonary fibrosis. Thorax. ISSN 00406376. 57, 3 (2002), 263–266.

SCHMITTGEN, T. D., LIVAK, K. J. - Analyzing real-time PCR data by the comparative C(T) method. Nature protocols. ISSN 1754-2189. 3, 6 (2008), 1101–1108.

SCHUCHERT, M. J., LUKETICH, J. D. - Solitary sites of metastatic disease in non-small cell lung cancer. Current treatment options in oncology. ISSN 1527-2729. 4, 1 (2003), 65–79.

SCHULTZ, M. J., SWINDALL, A. F., BELLIS, S. L. - Regulation of the metastatic cell phenotype by sialylated glycans. Cancer and Metastasis Reviews. ISSN 01677659. 31, 3-4 (2012), 501–518.

SPIRO, R. G. - Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. Glycobiology. ISSN 0959-6658. 12, 4 (2002), 43R–56R.

STAMPER, H. B., WOODRUFF, J. J. - Lymphocyte homing into lymph nodes: in vitro demonstration of the selective affinity of recirculating lymphocytes for high-endothelial venules. The Journal of experimental medicine. ISSN 0022-1007. 144, 3 (1976), 828–833.

STENBYGAARD, L. E. [et al.] - Metastatic pattern in non-resectable non-small cell lung cancer. Acta oncologica (Stockholm, Sweden). ISSN 0284186X. 38, 8 (1999), 993–998.

STOWELL, S. R., JU, T., CUMMINGS, R. D. - Protein Glycosylation in Cancer. Annual Review of Pathology: Mechanisms of Disease. ISSN 1553-4006. 10, (2015), 25-50.

STRELL, C., ENTSCHLADEN, F. - Extravasation of leukocytes in comparison to tumor cells. Cell Commun Signal. ISSN 1478-811X. 6, 10 (2008), 5.

TARP, M. A., CLAUSEN, H. - Mucin-type O-glycosylation and its potential use in drug and vaccine development. Biochimica et Biophysica Acta - General Subjects. ISSN 03044165. 1780, 3 (2008), 546–563.

TATENO, H. [et al.] - A novel strategy for mammalian cell surface glycome profiling using lectin microarray. Glycobiology. ISSN 09596658. 17, 10 (2007), 1138–1146.

TAYLOR, M. E., DRICKAMER, K. - Introduction to glycobiology [Online]. 3rd ed. Oxford New York: Oxford university press, 2011. [Consult. September 14 2015]. Available from WWW:<https://books.google.pt/books/about/Introduction_to_Glycobiology.html?id=4Y5wQgAACAAJ&redir_esc=y >. ISBN 978-0-19-956911-3.

TAYLOR-PAPADIMITRIOU, J. [et al.] - MUC1 and cancer. Biochimica et Biophysica Acta - Molecular Basis of Disease. ISSN 09254439. 1455, 2-3 (1999), 301–313.

TEH, E., BELCHER, E. - Lung cancer: diagnosis, staging and treatment. Surgery (Oxford). ISSN 02639319. 32, 5 (2014), 242–248.

THOMAS, S. N. [et al.] - Carcinoembryonic antigen and CD44 variant isoforms cooperate to mediate colon carcinoma cell adhesion to E- and L-selectin in shear flow. Journal of Biological Chemistry. ISSN 00219258. 283, 23 (2008), 15647–15655.

TIAN, E., HAGEN, K. G. T. - Recent insights into the biological roles of mucin-type O-glycosylation. Glycoconjugate journal. ISSN 02820080. 26, 3 (2009), 325–334.

TOGAYACHI, A. [et al.] - Up-regulation of Lewis enzyme (Fuc-TIII) and plasma-type α1, 3Fucosyltransferase (Fuc-TVI) expression determines the augmented expression of sialyl Lewis x antigen in non-small cell lung cancer. International journal of cancer. ISSN 00207136. 83, 1 (1999), 70–79.

TOZAWA, K. [et al.] - Positive correlation between sialyl Lewis X expression and pathologic findings in renal cell carcinoma. Kidney international. ISSN 00852538. 67, 4 (2005), 1391–1396.

UGORSKI, M., LASKOWSKA, A. - Sialyl Lewis a: a tumor-associated carbohydrate antigen involved in adhesion and metastatic potential of cancer cells. ACTA BIOCHIMICA POLONICA-ENGLISH EDITION-. ISSN 0001527X. 49, 2 (2002), 303–312.

VAJARIA, B. N. [et al.] - Expression of Glycosyltransferases; ST3GAL1, FUT3, FUT5, and FUT6 Transcripts in Oral Cancer. Glycobiology Insights. ISSN 1179-2515. 4, (2014), 7-14.

VARKI, A. [et al.] - Essentials of Glycobiology [Online]. 2nd ed. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press, 2009 [Consult. September 1 2015]. Available from WWW:< http://www.ncbi.nlm.nih.gov/books/NBK1908/>. ISBN 9780879697709.

VESTWEBER, D., BLANKS, J. E. - Mechanisms that regulate the function of the selectins and their ligands. Physiological reviews. ISSN 0031-9333. 79, 1 (1999), 181–213.

WANG, J. [et al.] - Expression and prognostic relevance of tumor carcinoembryonic antigen in stage IB non-small cell lung cancer. Journal of thoracic disease. ISSN 2077-6624. 4, 5 (2012).

WITZ, I. P. - The selectin-selectin ligand axis in tumor progression. Cancer and Metastasis Reviews. ISSN 01677659. 27, 1 (2008), 19–30.

WOLFERT, M. A., BOONS, G-J. - Adaptive immune activation: glycosylation does matter. Nature chemical biology. ISSN 1552-4469. 9, 12 (2013), 776–84.

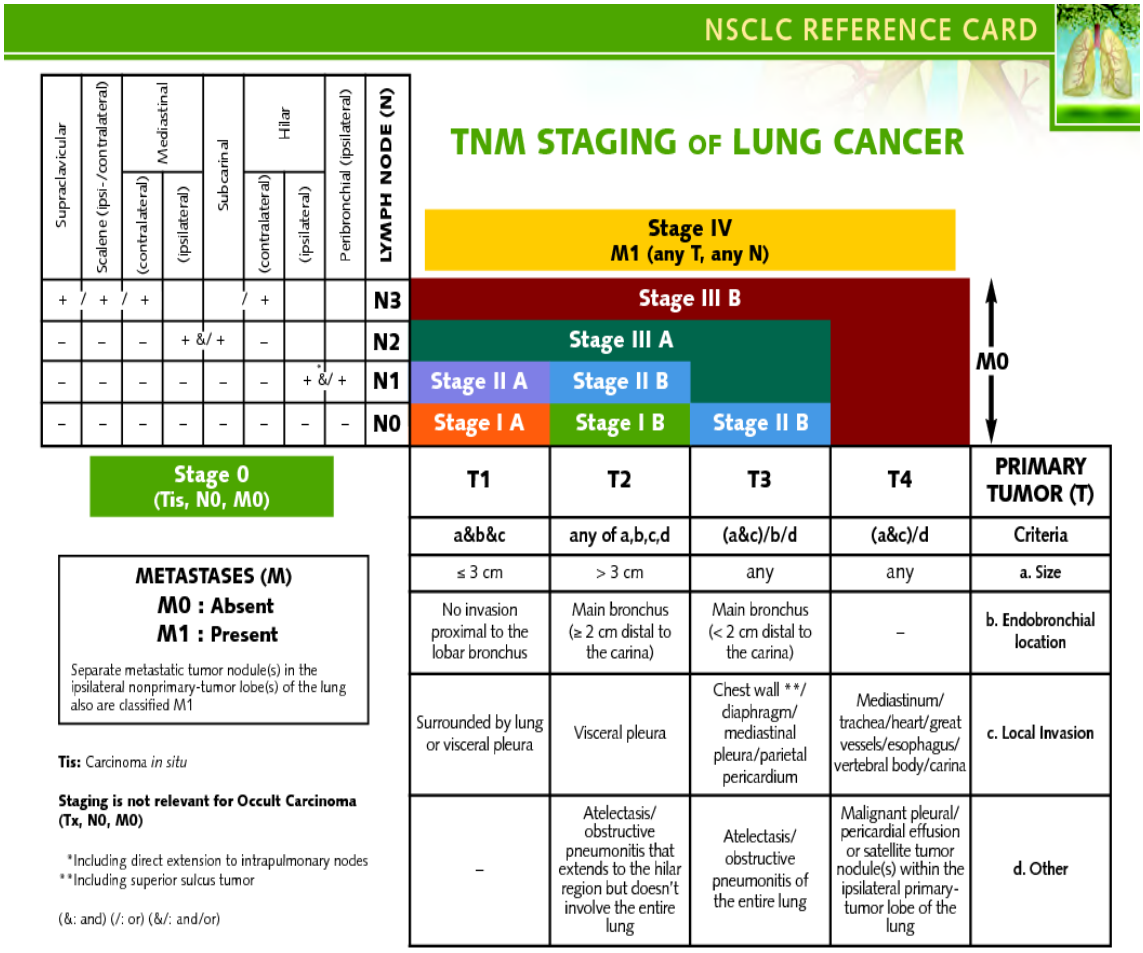
WONG, M. L., MEDRANO, J. F. - Real-time PCR for mRNA quantitation. BioTechniques. ISSN 07366205. 39, 1 (2005), 75–85.

YAREMA, K. J., BERTOZZI, C. R. - Characterizing glycosylation pathways. Genome Biol. ISSN 1465-6914. 2, 5 (2001).

6. Appendixes

Appendix 6.1 - TNM staging of Lung Cancer - 7th edition of TNM staging based on International Association for the Study of Lung Cancer.

Adapted from WWW:<http://pulmccm.org/main/2015/interventional-pulmonology/who-needs-mediastinoscopy-after-negative-ebus-staging-for-lung-cancer/.



Appendix 6.2 - Constitution of buffer solutions used in this work.

- **Phosphate buffered saline 1× (PBS 1×):** Solution with 1.47mM KH_2PO_4, 4.29mM $NaHPO_4.7H_2O$, 137mM NaCl e 2.6mM KCL (pH=7.4) in distilled water.

- **Lysis buffer:** Solution with 150mM NaCl, 2mM $CaCl_2$, 50mM Tris pH 7,4, 20 µg/mL PMSF (Sigma-Aldrich), 2% NP-40, one tablet of EDTA-free protease inhibitor cocktail (Roche) per 200mL and distilled water. Pre-cooling the solution in ice for 20min, before adding protease inhibitors.

- **Resolving Gel (8% - for two gels):** 240µL SDS 10%, 24µL TEMED (Sigma-Aldrich), 240µL APS, 11.3mL distilled water, 6mL resolving buffer (Tris 1.5M pH 8.8) and 6.3mL 30% Acrylamide/Bis Solution 37.5:1 (Bio-Rad).

- **Stacking Gel (for two gels):** 100µL SDS 10%, 10µL TEMED (Sigma-Aldrich), 100µL APS, 6mL distilled water, 2.52mL stacking buffer (Tris 0.5M pH 6.8) and 1.32mL 30% Acrylamide/Bis Solution 37.5:1 (Bio-Rad).

- **Running buffer (10×):** 30.39g Tris-base (Sigma-Aldrich), 144.1g glycine (Sigma-Aldrich) and 10g SDS (Sigma-Aldrich) in distilled water (1L).

- **Transfer buffer (10×):** 30.38g Tris-base (Sigma-Aldrich) and 144.1g glycine (Sigma-Aldrich) in distilled water (1L).

- **Transfer buffer (1×) -** 700mL distilled water, 100µL transfer buffer 10× and 200µL methanol (VWR) for 1L.

- **Tris-buffered saline (TBS 10×):** Solution with 20mM Tris-HCl and 150mM NaCl in distilled water (pH=7.6).

- **Recombinant Mouse E-selectin/CD62E Fc Chimera:** Reconstitute at 0.5 mg/mL in sterile, distilled water.

Appendix 6.3 - Patient's characteristics: Gender (F - female or M - male), Smoker (0 - Non-Smoker or 1- Smoker). Patients between number 1 and 21 were analyzed by RT-PCR. Patients between 22 and 33 were excluded from this study.

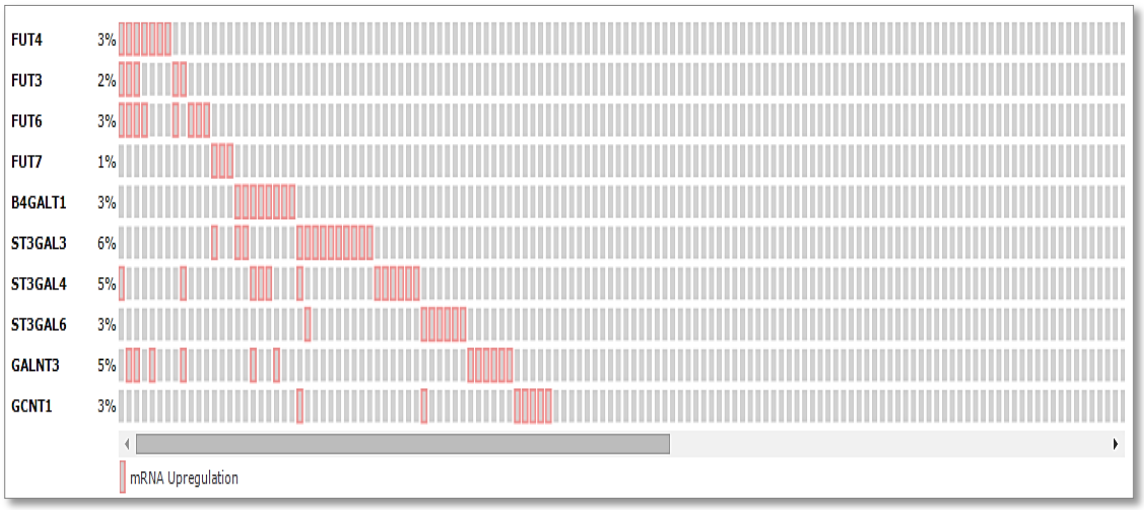
Patient identification	Gender	Age	Diagnostic	Stage	Smoker
1	F	83	AC	IA	0
2	M	66	AC	IB	1
3	M	67	AC	IA	1
4	M	60	AC	IIIA	1
5	M	80	SCC	IA	1
6	F	59	SCC	IA	1
7	M	65	AC	IB	1
8	M	64	SCC	IB	1
9	F	67	AC	IIIA	0
10	M	68	AC	IA	1
11	M	63	AC	IIIA	1
12	M	65	AC	IIA	1
13	F	64	AC	IIIA	0
14	F	71	AC	IA	0
15	M	68	AC	IIIA	1
16	M	52	AC	IIIA	0
17	M	49	SCC	IIIA	1
18	M	48	AC	IIB	0
19	M	49	AC	IA	0
20	M	72	AC	IIA	1
21	M	61	AC	IIIA	1
22	M	53	AC	IA	1
23	F	72	AC	IIIA	1
24	M	76	AC	IIIA	1
25	M	77	AC	IIB	1
26	F	67	AC	IB	0
27	M	72	AC	IA	1
28	M	78	AC	IIIA	0
29	M	84	AC	IA	1
30	M	64	AC	IA	1
31	F	55	AC	IA	1
32	M	69	AC	IIIA	1
33	M	81	AC	IA	1

Appendix 6.4 - Data sets from cBioPortal database.

Database website: http://www.cbioportal.org/

1st Study- Query used: 1- Lung Adenocarcinoma (TCGA, Nature 2014), 2 - mRNA Expression z-scores (RNA Seq V2 RSEM), 3 - Tumors with mRNA data (RNA Seq V2) (230), 4- FUT3, FUT4, FUT6, FUT7, ST3GAL3, ST3GAL4, ST3GAL6, B4GALT1, GCNT1, GALNT3

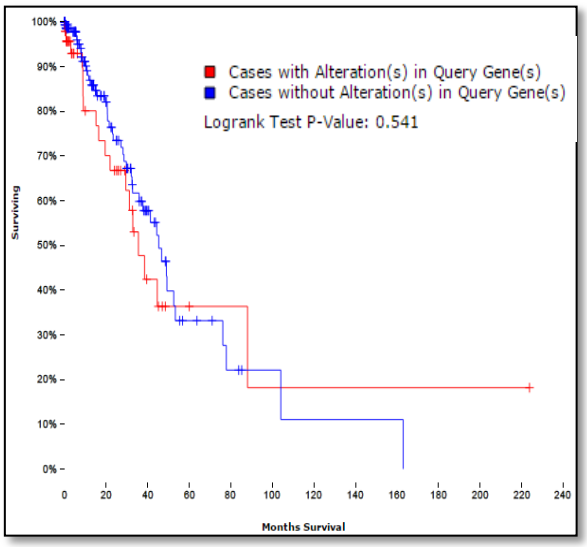
a) OncoPrit tab



b) Mutual exclusivity and co-occurrence analysis: p-value was determined by a Fisher's exact test (p-value < 0.05: significant association); Log Odds ratio quantifies how strongly alterations in gene A are associated with alterations in gene B; When Log Odds Ratio > 0, association towards co-occurrence and when Log Odds Ratio ≤ 0, association towards mutual exclusivity. Here, we only show the gene pairs with a significant association.

Gene A	Gene B	p-Value	Log Odds Ratio	Association	
FUT3	FUT4	<0.001	>3	Tendency towards co-occurrence	Significant
FUT3	FUT6	<0.001	>3	Tendency towards co-occurrence	Significant
FUT4	FUT6	<0.001	>3	Tendency towards co-occurrence	Significant
FUT3	GALNT3	0.001	>3	Tendency towards co-occurrence	Significant
FUT4	GALNT3	0.003	2.881	Tendency towards co-occurrence	Significant
ST3GAL4	B4GALT1	0.005	2.653	Tendency towards co-occurrence	Significant
FUT3	ST3GAL4	0.023	2.663	Tendency towards co-occurrence	Significant

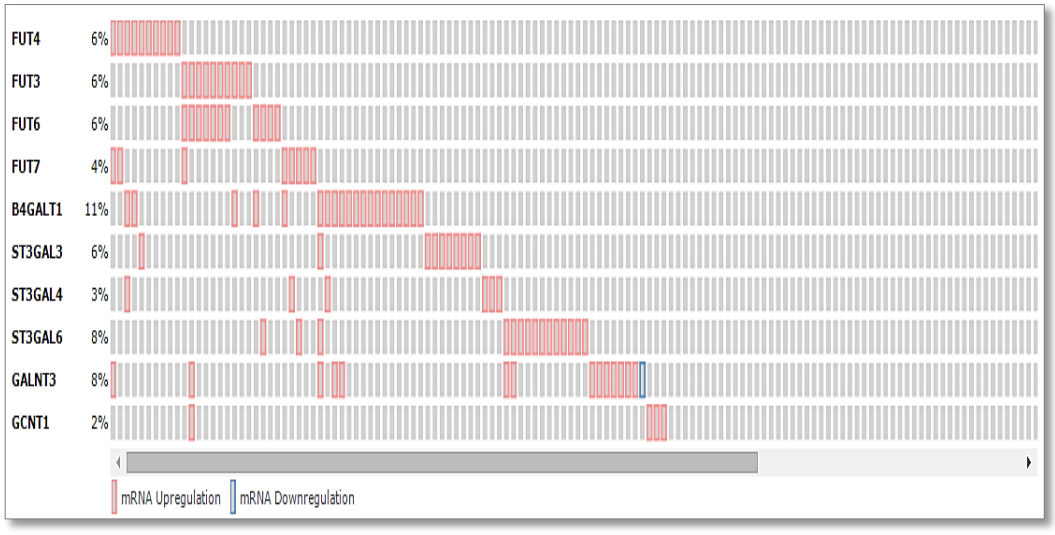
c) Overall Survival Kaplan-Meier Estimate: red curve includes all samples with alterations in one of selected genes and blue curve includes all samples without any alterations.



	#total cases	#cases deceased	median months survival
Cases with Alteration(s) in Query Gene(s)	49	18	35.52
Cases without Alteration(s) in Query Gene(s)	154	45	45.31

2nd Study - Query used: 1- Lung Squamous Cell Carcinoma (TCGA, Nature 2012), 2 - mRNA Expression z-scores (RNA Seq RPKM), 3 - Tumors with mRNA data (RNA Seq) (178), 4- FUT3, FUT4, FUT6, FUT7, ST3GAL3, ST3GAL4, ST3GAL6, B4GALT1, GCNT1, GALNT3

a) OncoPrit tab



b) Mutual exclusivity and co-ocurrence analysis

Gene A	Gene B	p-Value	Log Odds Ratio	Association
FUT3	FUT6	<0.001	>3	Tendency towards co-occurrence Significant

Note: For this study, there are not results about survival analysis.

Appendix 6.5 - Statistical analysis of Spearman correlation for normal samples (yellow cells) and tumor samples (orange cells). The values inside cells correspond to correlation coefficients and the correlation is significant at 0.05 (*), 0.01 (**), 0.001 (***) and 0.0001 (****) level.

	FUT3	FUT4	FUT6	FUT7	ST3GAL3	ST3GAL4	ST3GAL6	B4GALT1	GCNT1	GALNT3
FUT3		-0.125	0.453*	0.157	0.181	0.242	0.324	0.595**	0.251	0.465*
FUT4	0.305		0.021	0.429	0.452*	0.557**	0.159	-0.012	0.596**	0.474*
FUT6	0.558**	0.149		0.383	0.149	0.129	0.186	0.307	0.22	0.374
FUT7	0.614**	0.419	0.245		0.469*	0.615**	0.664**	0.223	0.64**	0.573**
ST3GAL3	0.543*	0.57**	0.217	0.644**		0.879****	0.757****	0.344	0.899****	0.809****
ST3GAL4	0.584**	0.358	0.11	0.713***	0.667***		0.758****	0.38	0.961****	0.874****
ST3GAL6	0.637**	0.369	0.31	0.75****	0.695***	0.771****		0.314	0.761****	0.653**
B4GALT1	0.827****	0.423	0.322	0.78****	0.528*	0.706***	0.723***		0.294	0.517*
GCNT1	0.709***	0.535*	0.457*	0.761****	0.67***	0.658**	0.865****	0.791****		0.916****
GALNT3	0.731***	0.537*	0.4	0.75****	0.874****	0.776****	0.854****	0.673***	0.849****	

Normal
Tumor

Appendix 6.6 - Values of NSCLC lysates quantification by Pierce® BCA Protein Assay Kit (Thermo Scientific).

Patient identification	**[protein in normal tissue] in µg/µL**	**[protein in tumor tissue] in µg/µL**
1	9.53	27.93
2	16.01	14.79
3	18.92	19.20
4	12.51	14.17
5	21.94	25.09
6	26.34	11.09
7	17.06	16.56
8	17.36	22.62
9	16.90	13.33
10	12.43	17.08
11	35.83	14.77
12	19.7	34.8
13	20.02	29.2
14	10.18	12.11
16	12.24	28.3
17	10.62	27.6
18	22.54	11.69
19	12.51	20.79

Appendix 6.7 - Poster presented at iMED 7.0 Conference Lisbon 2015 within Fundação AstraZeneca Innovate Competition.

Role of fucosyltransferase 3 in hematogenous metastasis of lung cancer

I. Ferreira[1], M. Carrascal[1], A. Bugalho [1,2], A. Fernandes[3], P. Baptista[3] and P. Videira[1,3]

Contact information: ines.ferreira@fcm.unl.pt

1 - CEDOC, Chronic Diseases Research Center, NOVA Medical School/Faculdade de Ciências Médicas, Universidade Nova de Lisboa, Campo dos Mártires da Pátria, 130, 1169-056 Lisbon, Portugal

2 - Hospital CUF Infante Santo e Hospital CUF Descobertas, Lisbon, Portugal

3- Departamento Ciências da Vida, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Lisbon, Portugal

Introduction

Lung cancer (LC) is the leading cause of cancer death around the world, caused mainly by the metastatic spread of primary tumor cells. Hematogenous metastasis of LC still poorly understood and represents the life threatening event in this malignancy. Thus, it is essential to understand and identify mechanisms associated with the development of tumor cell invasion and metastasis.

The interaction between circulating cancer cells and the vascular endothelium is one of the most critical steps for transmigration and metastasis. LC is known to overexpress endothelial selectins ligands (E-SL), prototypically the glycans sialyl Lewis^x (sLex) or sialyl Lewis^a (sLea). However, its functional role in allowing LC adhesion to endothelium is still poorly understood. Here, we intended to contribute to a better understanding of the molecular effectors that mediate LC interaction with endothelium and further influence in tumor progression. A better understanding of the carriers for E-SL will allow the identification of tumor specific targets for novel therapies.

Our work aims to :

- Study the expression of the genes involved in the biosynthesis of sLex and sLea, with a special focus on 1,3/4-fucosyltransferase (FUT's), crucial enzymes that catalyze the transfer of fucose residue to glycoproteins/glycolipids;
- Analyze the expression of E-SL in normal and tumor LC tissues;
- Associate E-SL expression on tumor cells with a better ability to bind to endothelial selectins (E-selectins) under flow conditions.

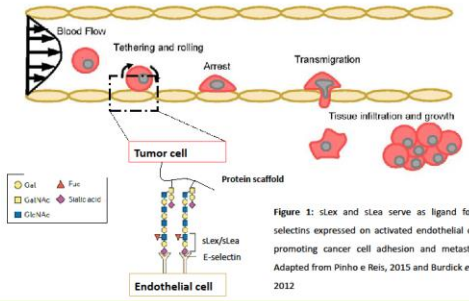


Figure 1: sLex and sLea serve as ligand for E-selectins expressed on activated endothelial cells, promoting cancer cell adhesion and metastasis. Adapted from Pinho e Reis, 2015 and Burdick *et al.*, 2012

Results

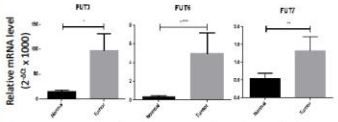


Figure 2: Gene expression values of FUT3, FUT4 and FUT7 in normal and tumor samples of LC patients analyzed by Real Time-PCR. Values infer the number of mRNA molecules of a certain gene per 1000 molecules of the average of the endogenous controls (β-actin and GAPDH). p<0.05 (*), p<0.01 (**), p<0.0001 (****)

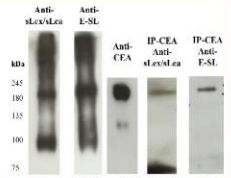


Figure 3: Western blotting analysis of total tumor lysate and carcinoembryonic antigen immunoprecipitate (IP-CEA) of patient 7, stained for sLex/sLea, all E-SL and CEA.

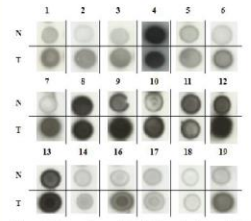


Figure 4: Expression of E-SL in patient's matched normal/tumor tissue pairs. Matched tumor (T) and normal samples (N) from LC patients were homogenized, dot blotted and stained with anti-sLex/sLea. The numbers correspond to the number of the patient.

A

		LC cell lines		
		H292	H1299	A549
Gene expression values	FUT3	64.576	0.104	0.058
	FUT4	0.179	0.380	0.641
	FUT6	0.042	0.001	0.000
	FUT7	0.001	0.005	0.001
MFI values	sLex/sLea	5706	102	24
	E-SL	38639	10201	4775

B

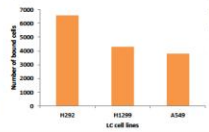


Figure 5:
A- Gene expression values of FUT's and Median Intensity Fluorescence (MFI) values for sLex/sLea and E-SL in three LC cell lines.
B – Adhesion of LC cancer cells to E-selectin. Data shows normalized values corresponding to the number of bound cells.

Conclusions

- **Overexpression of FUT 3, 6 and 7** in LC tumor tissues
- **↑ expression of E-SL** in LC tumor samples
- **CEA as an E-selectin ligand** decorated **with sLex/sLea** in LC patients
- LC cell lines → **Overexpression of FUT3 → ↑ E-SL → Increased ability to adhere to E-selectin**

FUT's, especially FUT3, are important targets to modulate E-SL expression

Future Perspectives

Inhibition of all FUT's expression with a **fluorinated fucose**

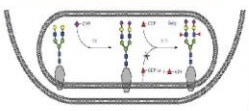


Figure 6: Fluorinated monosaccharide analogs act as metabolic glycosyltransferase inhibitors. Adapted from Rillahan et al., 2012

Inhibition of FUT3 with a RNAi approach (e.g. siRNA and antisense)

Inhibit FUT3 → **↓ E-SL expression** → **↓Metastatic potency of LC cells**

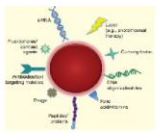


Figure 7: Multifunctional gold nanoparticles represent powerful theranostic platforms. Adapted from Cabral and Baptista, 2014

Translational application of the FUT3 inhibition in patients

Specific therapy with Nanoparticles

Methods

- **Normal and tumor non-small cell lung cancer (NSCLC) tissues**
  - Gene expression by **Real Time-PCR**
  - Expression of sLex/sLea by **dot blotting**
  - Identification of E-SL by **western blotting**
- **LC cell lines (H1299, H292 and A549)**
  - Gene expression by **Real Time-PCR**
  - Phenotypic analysis by **flow cytometry**
  - **Adhesion assays** to E-selectin

Genes analyzed: FUT3, FUT4, FUT6, FUT7, ST3GAL3, ST3GAL4, ST3GAL6, GCNT1, β4GALT1, GALNT3 and CEA

sLex/sLea and E-SL expression & Protein scaffolds of E-SL : CEA, MUC1, CD44, MUC5AC, MUC5B

References

- Burdick, M.M., Henson, K.A., Delgadillo, L.F., Choi, Y.E., Goetz, D.J., Tees, D.F.J., and Benencia, F. (2012). Expression of E-selectin ligands on circulating tumor cells: cross-regulation with cancer stem cell regulatory pathways? Front. Oncol. *2*.
- Cabral, R.M., and Baptista, P. V (2014). Anti-cancer precision theranostics: a focus on multifunctional gold nanoparticles. Expert Rev. Mol. Diagn. *14*, 1041–1052.
- Häuselmann, I., and Borsig, L. (2014). Altered tumor-cell glycosylation promotes metastasis. Front. Oncol. *4*.
- Pinho, S.S., and Reis, C. a. (2015). Glycosylation in cancer: mechanisms and clinical implications. Nat. Rev. Cancer *15*, 540–555.
- Rillahan, C.D., Antonopoulos, A., Lefort, C.T., Sonon, R., Azadi, P., Ley, K., Dell, A., Haslam, S.M., and Paulson, J.C. (2012). Global metabolic inhibitors of sialyl-and fucosyltransferases remodel the glycome. Nat. Chem. Biol. *8*, 661–668.